

HIGH FREQUENCY PRICE DECOMPOSITION AND TRADE ARRIVALS

ALEXANDER R. KENT

Advisor: Carlos Martins-Filho (Economics)
Honors Representative: Martin Boileau (Economics)
Outside Member: Daniel Brown (Finance)

April 6, 2015

Department of Economics
University of Colorado
Boulder, CO 80309-0256
Email: alexander.r.kent@colorado.edu

Abstract¹. Typically, the prices of financial assets are studied over fixed-time intervals such as the case with monthly or daily returns. Modern technology now allows us to consider each transaction that occurred throughout a trading period and the particular instance in time at which it was placed. Statistical analysis of financial assets conducted at this level is referred to as high frequency econometrics. This microscopic view of the market allows us to observe an asset's price formation process in continuous time. High frequency data is marked by a number of peculiarities that do not persist in discrete-time financial data, thus requiring a different econometric approach in order to preserve the vast amount of microstructure information embedded in the transaction data. In this paper, we construct and specify the joint probability distribution of price movements and trade arrivals as a compound Poisson process to build a theoretical framework to study the interplay of volatility and the timing of trades. We extend the price decomposition model proposed by Rydberg and Shephard (2003) by defining the magnitude of price change process to follow an adaptation of the autoregressive conditional multinomial—a finite state, VARMA model originally developed by Engle and Russell (2005). Furthermore, we define the trade arrival process to be a doubly-stochastic Poisson process (or Cox process) and propose estimating its random intensity through kernel density estimation.

Keywords: high frequency econometrics, transaction prices, trade arrivals, market microstructure

¹This paper was written as an Undergraduate Honors Thesis in Economics at the University of Colorado - Boulder.

1 Introduction

Historically, financial markets have been a felicitous area for econometric research due to the abundance of directly observable data that is readily available in relation to other economic systems of interest. A popular approach in analyzing financial assets has been to consider a sequence of prices, or returns, recorded at fixed time intervals; for example, S&P 500 yearly, monthly, and daily returns. These returns are calculated as the change in the last settled price of the asset over the particular time interval. Fixed interval approaches for modeling asset returns are advantageous in that modeling the time series in discrete time gives us access to a rich, existing econometric toolbox. However, this discretization ignores a plurality of the trading mechanisms and market dynamics which determine how an asset achieves its price. Consequently, an approach of this kind is insufficient to supply a complete, accurate description of the equity's price formation process. Modern technology and improved data management now allows us to observe every transaction recorded and the particular instance in time at which it was placed throughout the trading day. Datasets of this nature provide us with an unprecedented view of trading at an infinitesimal level. Accordingly, we are no longer considering closing prices consolidated across all exchanges over an aggregated interval, but rather the particular price agreed upon between matched market participants at an individual exchange. Literature has referred to the financial time series observed at this granularity as high frequency data.

High frequency data possesses many unique characteristics that are not found in other financial time series due to the asset price's sensitivity to the particular set of rules governing the mechanics of trading. In a market such as the New York Stock Exchange (NYSE) or NASDAQ, orders can arrive at any instance causing the trades to be irregularly spaced throughout time. Furthermore, institutional rules require exchanges to maintain a minimum unit of price increments known as ticks forcing transaction prices to live on a discrete grid. Consequently, in order to preserve the vast microstructure information embedded in high frequency data we must adapt an econometric

approach different from those typically employed to analyze financial assets. This paper is concerned with developing statistical models that can capture the behavior of equities at the trade-by-trade level in continuous time. Motivated by questions regarding how prices evolve over the trading day and their interaction with other microstructure variables, we construct models for trading price, volatility, and transaction arrival rate. The methodological framework we provide can be implemented to affirm prevailing market microstructure theory. Easley and O'Hara (1992) contend that a high trading intensity is likely a strong indication of the presence of informed traders. In such a situation, the market specialist commonly will increase the price's sensitivity to the order flow which induces higher volatility. Additionally, Diamond and Verrecchia (1987) suggest that negative information cannot be incorporated as quickly into a stock's price due to specific constraints on short selling. If this is indeed the case then slow trading rates should be closely associated with bad news and falling stock prices, while high trading rates should indicate good news and rising stock prices.

We begin by defining the economic variables of interest and representing the problem probabilistically. Let Z_i denote the price change of an asset resulting from the i^{th} trade. Each trade occurs at a random point in time generated by a stochastic point process. The primary objectives of this work are to construct and estimate the joint distribution of price movements and the stochastic point process describing the arrival of trades. In a seminal paper by Rydberg and Shephard (2003), to reveal additional trade information not previously apparent, the researchers propose decomposing Z_i into a product of three component variables: activity, direction, size. They chose to estimate probability distribution of price movement activity and direction through auto-logistic regression and size by a negative binomial generalized linear model, which was chosen for its simplicity and familiarity. In this paper, we extend the price decomposition model by estimating the size process with an autoregressive conditional multinomial (ACM) model—a continuous-time, discrete-state process originally developed by Engle and Russell (2005). Although alluding to potential approaches, Rydberg et al. are unspecific about the structure of the trade arrival process which excites the Z_i . Furthermore, we define the trade arrival process to be a doubly-stochastic Poisson process (or Cox

process) and propose determining the trade arrival intensity process via kernel density estimation due to the estimator's ability to continually learn from the data and to provide insight toward the specification of a more descriptive model in the future.

This paper proceeds as follows. Section 2 details the dataset used in this research and the unique characteristics of the trade data. Section 3 supplies an overview of the high frequency literature. Sections 4 and 5 will cover the methodological framework for modeling high frequency asset returns and transaction arrival rates. In Section 6 we conclude and offer possible directions for subsequent research.

2 Data

One of the most predominant dissimilarities between high frequency financial econometrics and low frequency (e.g., monthly or daily time intervals) financial econometrics is the nature in which the data studied is formed and accumulated. At low frequencies, typically one is concerned with the changes in price of an asset calculated over a particular holding period. For example, the analysis of IBM daily returns involves a series of prices calculated on a continuous scale indexed by a specific fixed time interval. Here, the prices used are the last settled price at the end of that trading day.

However in high frequency financial data, these characteristics do not persist. Rather they are marked by a number of fundamental peculiarities. Since we are considering asset prices at the transaction level, returns can no longer be considered at fixed time periods as trades arrive at random points in time. Furthermore, the prices themselves that we are observing are inherently different from those observed at lower frequencies. The price recorded is the price that a particular pair of market participants agreed upon to trade a specific amount of the asset. In modern practice, many buyers and sellers are connected by matching algorithms which prioritize order selection by finding the National Best Bid-Offer¹ (NBBO) price which is defined to be the lowest ask/offer (what a dealer is willing to sell at) and the highest bid (what an investor is willing to pay) quotes

¹This price is established and enforced by the SEC's Regulation NMS, which was enacted in 2005.

available on all exchanges. This price is shown to the public through the Securities Information Processor (SIP), which links all U.S. exchanges and consolidates protected quote information and disseminates it to display regulatory information, like the NBBO. Consequently, through this data we observe the price innovation process as arriving trades impute market information and investor sentiment helping the asset achieve an equilibrium-trading price in continuous time. An additional microstructure feature of this price data not found in low frequency data is that prices are restricted to take on discrete integer amounts, known as ticks, while over longer periods of time security prices appear to be continuously valued random variables. This perception exists since the asset's price volatility exceeds the effects of the restricted discrete price changes as time goes on, diminishing the bias associated with treating price as continuous. Presently, the value of a tick on U.S. exchanges as designated by the Securities Exchange Commission (SEC) is equal to \$0.01²

In developing our high frequency models, we referenced 12 months of NYSE Euronext trade and quote data (TAQ) spanning from April 2010 to March 2011 for Bank of America (BAC) and Abbott Laboratories³ (ABT) on the New York Stock Exchange (NYSE). Our dataset contains trade information for every transaction that occurred during this time period for these equities and the instance in time that it was recorded down to the second. Each data point displays the date of the trade, the transaction price, the timestamp, and the number of shares exchanged (volume). In practice, traders have the ability to place orders at the millisecond level via low latency data connections meaning the accuracy of our measurements do not perfectly define the trade arrival times. As a result there are some instances where multiple trades occurred with identical timestamps. Any further assumptions about the ordering of these trades would be *a priori*, inducing unnecessary bias, restricting our ability to analyze the price innovation process. To handle this problem, we maintain a one-to-one relationship between transaction prices and trade times. Empirically, we accomplish this by following the suggestion of Jasiak and Gouriéroux (2001), computing a weighted

²The tick size was stipulated in Reg. NMS, "to limit the ability of a market participant to gain execution priority over a competing limit order by stepping ahead by an economically insignificant amount".

³The trade data for ABT predates the company's October 2011 separation into two publicly traded entities: Abbott Labs, specializing in medical products; AbbVie, specializing in pharmaceutical research.

average of the trade price and volume at each instance where more than one trade was recorded at a specific second. While preserving information about the arrival times, this transformation causes the price increments to no longer take on integer tick values, but rather continuous values. Modeling security prices in tick time continuously is unfavorable since many of the prices we observe in this new series are unobtainable in practice. Moreover, the price movements relative to the trading price of the security at the high frequency level are generally quite small. This means that a realistic high frequency model must be able to highlight the asset price's sensitivity to the order flow, which is best accomplished by means of a discrete-state model. To retain the discreteness of our price series, we implement a rounding procedure discussed by Engle and Russell in Ait-Sahalia's *Handbook of Financial Econometrics* (2010).

Another trading phenomenon that induces additional microstructure noise in the dataset arises from the natural discontinuity between consecutive trading days. Simply concatenating each daily series is insufficient as it neglects important trading mechanics that are specific to the particular time and day of the week. Orders, which are placed outside of trading hours, are filled through a call auction at the beginning of the next trading day. Moreover, a significant amount of trading occurs during the market's closing hour as traders close their daily positions and prepare for the subsequent trading day. The price formation processes at these times are distinct from other hours of the day and are a fundamental features of the trading dynamic at the NYSE. Rydberg et al. propose truncating the first and last 30 minutes of trading in order to eliminate the residual effects of the call auction and unusual high volatility near closing before amalgamating the time series (2003). Although a viable approach, this method causes us to ignore the two most active and potentially informative⁴ periods of the trading day. As a result, we propose grouping the trade information by day to observe the entirety of the market's operation hours and to allow for heterogeneity among model parameters. This approach is preferable as it would be reasonable to expect that the trading process may not only be dependent on the time of day, but on the specific day of the week, week of

⁴This assumption follows from the work of Easley and O'Hara (1982), who contend that high trading intensity is an indication of the presence of informed traders.

the month, and month of the year as well.

3 Literature Review

Methods frequently employed in financial econometrics often rely on the discrete indexing of fixed time intervals for financial data. This is evident in approaches such as in the ARCH framework for analyzing daily stock returns and volatility described in Poon and Granger (2003). However, high-frequency data possesses many unique features and irregularities that do not persist in lower frequency financial time series. As a result, much of the existing literature on high-frequency econometrics is dedicated toward constructing models that deal directly with these distinct characteristics such as the random spacing of trades, discrete price movements, and microstructure noise, rather than modifying the dataset to fit existing models.

Perhaps the most salient characteristic of high-frequency data is the irregular spacing of transactions through time (Engle 2000). Consequently, relying on popular, discrete-time models is insufficient for properly describing high-frequency time series without masking interesting microstructure features and inducing unnecessary bias (Engle, Russell 1998). Handling the erratic spacing of transactions requires the utilization of a stochastic point process; where in the application of modeling trade arrivals, is commonly termed a financial point process. Financial point processes can be constructed from two primary viewpoints—duration and intensity. Although similar in approach, the two can provide a different economic interpretation of the transaction arrival process. Duration is particularly useful in describing the likelihood of subsequent price changes [13] and the waiting time for new information. Intensity on the other hand, offers a more natural basis for measuring instantaneous volatility and is easily extended to the multivariate case, unlike duration [21]. Duration modeling of financial point processes involve predicting when a trade will occur, given a trade has not occurred since the last observation. Duration modeling of transaction times is strongly tied to the subject of survival analysis by considering the transaction duration to be the survival time. In a seminal paper by Engle and Russell (1998), they apply an autoregressive conditional duration

model, a variation of a dependent Poisson process, to estimate transaction duration. They formulate the duration similar to the proportional hazard model as originally proposed by Cox (1972) by decomposing the hazard to a product of the arrival density and a function of its covariates. Though other forms have been created, generally the ACD model is described by $\tau_i = t_i - t_{i-1}$, the time between trades and its conditional expectation $E[\tau_i|\mathcal{F}_{i-1}] = \psi_i$, such that

$$\psi_i = \omega + \sum_{j=0}^p \alpha_j \tau_{i-j} + \sum_{j=0}^q \beta_j \psi_{i-j} \quad (1)$$

By specifying the conditional intensity process as the hazard rate, conditional on all past information, the model provides a powerful framework for assessing the interaction between trade duration and asset volatility. Engle et al. (1998) find evidence to suggest that transactions are highly clustered due to the crowding of informed traders when the prevailing bid-ask spread is small.

Trade intensity models, which seek to measure the probability of observing a transaction at any point in time, follow a similar methodology to duration, but the variables of interest are inverses of each other. An example of an approach to intensity modeling is the use of Hawkes processes, as in Bowsher (2006), which specifies the intensity process as a self-exciting process driven by the time distance to past arrivals in the point process. A useful feature of Hawkes processes is that they can be fitted to handle clusters of arrivals, a phenomenon one would certainly expect in high-frequency finance such as when new information becomes available to traders who react nearly simultaneously. More usual forms of point processes do not have this feature and limited to only allowing one arrival at an instance in time. Following the work in trade duration modeling by Engle and Russell (1998), Hamilton and Jorda (2002) develop an intensity analog that corresponds to the inverse of the conditional duration. By doing so, they extend the ACD model of Engle and Russell to permit time-varying covariates. Zhang and Kou (2010) provide a strong framework for estimating arrival rates and autocorrelation functions associated with a Cox process by means of kernel density estimation. While applied in a biophysical context, the researchers contend that their methods developed can be applied seamlessly to other Cox processes exhibiting potentially both short-term and long-term

temporal dependence. We suspect this to be the case with high-frequency financial data given intuitive assumptions about intraday and seasonal trading patterns in the market. Furthermore, the non-parametrically derived autocorrelation function may be able to give us insight as to how trade intensity is distributed, aiding in the testing of future parametric models (Bauwens and Hautsch 2007).

A sometimes overlooked feature of financial high-frequency data is that the price changes are restricted to live on a discrete grid due to restrictions imposed by regulatory agencies. Trade-by-trade price movements are expressed in terms of an elementary value, dubbed a tick. This is contrary to low frequency pricing, in which assets appear to take on continuous price values due to smoothing implemented by market specialists. Thus the discretization of price becomes an important feature of this data and can be empirically complex to handle. In practice, we only observe a small collection of different tick-valued price changes. In their 2005 study, Engle and Russell find that 99.3% of all of their observed trades took on only one of five values, down 2 ticks to up 2 ticks. Engle and Russell extend their ACD model to jointly model price and duration in a method they call autoregressive conditional multinomial autoregressive conditional duration (ACM-ACD). The motivation for constructing the ACM model to describe high-frequency data is derived from similar models' success in handling highly temporally dependent data, such as those found in option pricing. A unique approach to simplify the construction of the joint price movement distribution is proposed in Rydberg and Shephard (2003) by decomposing the stochastic process into a product of conditional densities describing three fundamental features of the economic process. This decomposition allows the researchers to test the serial dependence of price movements on past activity, direction of change, and magnitude of change. Microstructure noise such as bid-ask bounce, which is prevalent in high-frequency data, can also be tested for under this framework. However, this model has many areas for which further research can improve upon. Through decomposition, information about stock dynamics becomes much more apparent, but inference on the model implemented empirically by the researchers is limited by distributional misspecification.

4 Price Movements

A defining feature of high frequency asset behavior is that while prices evolve through continuous time, the changes in the price or returns associated with each trade are restricted to live on a discrete grid. This fact arises due to policies maintained by the exchanges which specify minimum price increments, known as ticks, that securities can take on. Let us consider a general pricing model

$$p(t) = p(0) + \sum_{i=0}^{N(t)} Z_i \quad (2)$$

where $p(t)$ denotes the price of the asset at time $t \in \mathbb{R}$. Here, $N(t)$ denotes the number of trades realized between time 0 and time t and Z_i represents the price movement associated with the i^{th} trade. $N(t)$ acts as a counter, exciting Z_i at the arrival of each trade and is modeled by a family of stochastic processes referred to as financial point processes in the high frequency literature. This subject will be discussed in greater detail in Section 5. Because Z_i is restricted to exclusively take on multiples of the smallest price increment specified by the exchange it is being traded on, Z_i can be viewed as an integer process, which in practice takes only a handful of values. In this section, Z_i is modeled as being dependent only on itself, though this stipulation will be relaxed in subsequent sections to include information about trade arrivals. As a result, we have that $Z_i \in \mathbb{Z}$ and its natural filtration being $\mathcal{F}_i = \sigma(Z_j : j \leq i)$. Now we can formulate the joint probability distribution of the price movements as follows

$$P(Z_1, \dots, Z_n | \mathcal{F}_0) = \prod_{i=1}^n P(Z_i | \mathcal{F}_{i-1}), \quad (3)$$

by decomposing Z_i into a product of probabilities conditioned on all prior trade information. The principal motivation of this section is to construct and estimate this joint distribution. Directly, this can be an arduous task; however, following the suggestion made by Rydberg and Shephard (2003), we can simplify the process econometrically by decomposing Z_i into a product of three fundamental components—activity, direction, and size. Doing so allows us to further inspect the

determining factors and characteristics of price change such as asymmetrical returns and mean-reverting behavior. Moreover, under this framework we can better locate, and then control for, instances of microstructure noise such as bid-ask bounce which many models have not taken fully into account.

4.1 Decomposition

Through the preceding decomposition, we define the price movement corresponding to the i^{th} trade as

$$Z_i = A_i D_i S_i, \quad (4)$$

where A_i , D_i , S_i are defined to be activity, direction, and size, respectively. We define the activity series as a binary variable such that

$$A_i = \begin{cases} 1, & \text{if there is a price change from the } i^{th} \text{ trade} \\ 0, & \text{if there is no price change from the } i^{th} \text{ trade,} \end{cases} \quad (5)$$

the direction series conditioned on the i^{th} trade being active as a binary variable such that

$$D_i | (A_i = 1) = \begin{cases} 1, & \text{if the price change from the } i^{th} \text{ trade is positive} \\ -1, & \text{if the price change from the } i^{th} \text{ trade is negative,} \end{cases} \quad (6)$$

and the magnitude series conditioned on the i^{th} trade being active as an integer variable such that

$$S_i | (D_i, A_i = 1) = 1, 2, 3, \dots \quad (7)$$

Consequently, by Baye's Rule the distribution of price movements can be formulated as

$$P(Z_i | \mathcal{F}_{i-1}) = P(A_i D_i S_i | \mathcal{F}_{i-1}) = P(A_i | \mathcal{F}_{i-1}) P(D_i | A_i, \mathcal{F}_{i-1}) P(S_i | D_i, A_i, \mathcal{F}_{i-1}) \quad (8)$$

Note that $A_i = 0$ implies that $Z_i = 0$. While we could potentially model these series independently, let us not forget the motivation behind the decomposition, which remains constructing a multivariate model for the Z_i . We contend that imposing this configuration will yield richer, interpretable results not readily apparent when modeling Z_i directly.

4.2 Preliminary Component Models

Recall the price activity variable, A_i , which we pose as a binary variable indicating whether the i^{th} trade resulted in a non-zero price change. Since ultimately we are concerned with the role A_i has in determining Z_i , it suffices to examine the case where $A_i = 1$; otherwise, considering D_i and M_i is trivial. For this reason, we are interested in the behavior of the probability $p_i = P(A_i = 1 | \mathcal{F}_{i-1})$ over time. Initially, we assume that price activity obeys an auto-logistic structure such that

$$p_i = \frac{e^{\theta_i^A}}{1 + e^{\theta_i^A}} \quad \text{where} \quad \theta_i^A = \ln \left(\frac{p_i}{1 - p_i} \right) = \phi_0 + \boldsymbol{\phi} \mathbf{x}_i + \sum_{l=1}^L \beta_l A_{i-l}, \quad (9)$$

where ϕ_0 is a constant, $\boldsymbol{\phi}$ is a r -dimensional parameter vector, \mathbf{x}_i is a $r \times 1$ vector composed of potential elements of \mathcal{F}_{i-1} , β_l are parameters, and A_{i-l} are l -lag values of A_i . A logistic approach is appropriate since it allows us to extract the log odds of a trade producing a price change regressed on prior trade information (Cox, 1958). Additionally, since the direction of price change D_i (assuming $A_i = 1$) is also a binary variable, it takes on a similar structure where the probability of interest, $\delta_i = P(D_i = 1 | A_i = 1, \mathcal{F}_{i-1})$, is defined to be

$$\delta_i = \frac{e^{\theta_i^D}}{1 + e^{\theta_i^D}} \quad \text{where} \quad \theta_i^D = \ln \left(\frac{\delta_i}{1 - \delta_i} \right) = \kappa_0 + \boldsymbol{\kappa} \mathbf{y}_i + \sum_{l=1}^L \gamma_l D_{i-l}, \quad (10)$$

where κ_0 is a constant, $\boldsymbol{\kappa}$ is a r -dimensional parameter vector, \mathbf{y}_i is a $r \times 1$ vector composed of potential elements of \mathcal{F}_{i-1} , γ_l are parameters, and D_{i-l} are l -lag values of D_i . Since most traders are naturally risk-averse, in practice there tends to exist an asymmetrical response in volatility to up and downward price movements. To observe this phenomenon, we allow

$$S_i | (D_i, A_i = 1) \sim \begin{cases} g(\lambda_{ui}), & \text{if } D_i = 1, A_i = 1 \\ g(\lambda_{di}), & \text{if } D_i = -1, A_i = 1, \end{cases} \quad (11)$$

where $g(\lambda_{ki}) = P(S_i = s_i | D_i, A_i = 1) = \lambda_{ki} (1 - \lambda_{ki})^{s_i - 1}$ denotes the geometric probability distribution with parameter λ_{ki} , as proposed by Tsay (2010) which is a simplified version of the negative binomial GLARMA model implemented by Rydberg et al. (2003). The geometric parameter values evolve temporally as

$$\lambda_{ki} = \frac{e^{\theta_{ki}^S}}{1 + e^{\theta_{ki}^S}} \quad \text{where} \quad \theta_{ki}^S = \ln \left(\frac{\lambda_{ki}}{1 - \lambda_{ki}} \right) = \nu_{k0} + \boldsymbol{\nu}_k \mathbf{w}_{ki} + \sum_{l=1}^L \psi_{kl} S_{i-l}, \quad (12)$$

where $\nu_{k0}, \nu_k, \mathbf{w}_{ki}, \psi_{kl}$, and S_{i-l} play their logical roles. When considered in aggregate, the above models suggest that for the i^{th} trade Z_i exists in one of three states:

$$Z_i = \begin{cases} 0, & \text{if } A_i = 0, \text{ with probability } (1 - p_i) \\ g(\lambda_{ui}), & \text{if } A_i = 1, D_i = 1, \text{ with probability } p_i \delta_i \\ g(\lambda_{di}), & \text{if } A_i = 1, D_i = -1, \text{ with probability } p_i(1 - \delta_i) \end{cases} . \quad (13)$$

4.2.1 Estimation

By formulating Equations (2) and (7) in terms of the three states specified by our model, we obtain

$$\begin{aligned} P(Z_i = z_i | \mathcal{F}_{i-1}) &= \mathbb{1}_{1i}(1 - p_i) + \mathbb{1}_{2i}p_i\delta_i g(\lambda_{ui}) + \mathbb{1}_{3i}p_i(1 - \delta_i)g(\lambda_{di}) \\ &= \mathbb{1}_{1i}(1 - p_i) + \mathbb{1}_{2i}p_i\delta_i\lambda_{ui}(1 - \lambda_{ui})^{z_i-1} + \mathbb{1}_{3i}p_i(1 - \delta_i)\lambda_{di}(1 - \lambda_{di})^{z_i-1}, \end{aligned} \quad (14)$$

where $\mathbb{1}_{ji} = 1$ if the j^{th} state occurs, 0 otherwise. We now construct the log-likelihood function

$$\ln[P(Z_1 = z_1, \dots, Z_n = z_n | \mathcal{F}_0)] = \sum_{i=1}^n \ln[P(Z_i = z_i | \mathcal{F}_{i-1})], \quad (15)$$

to permit estimation of the parameters associated with the aggregate model mentioned above via maximum likelihood estimation.

4.3 Autoregressive Conditional Multinomial

In a similar spirit to Rydberg et al. (2003), Engle and Russell (2005) construct an autoregressive model for the conditional distribution of discrete price changes which they call the Autoregressive Conditional Multinomial (ACM) model. They begin by constructing a $k \times 1$ state vector, $\tilde{\mathbf{x}}_i$, whose elements indicate a particular integer increment of price change. A disadvantage of this approach is that number of ticks a stock can move is predetermined to be finite, whereas the Decomposition model permits a countably infinite number of tick moves. The impact of this tradeoff is diminished in practice, however, since the trading of most equities produces only a small collection of possible price changes. Based off of summary statistics of their data and to maintain parsimony, Engle and Russell choose $k = 5$ such that $\tilde{\mathbf{x}}_i$ indicates the occurrence of an element from the set of possible price changes $\Delta P_i = \{-2, -1, 0, 1, 2\}$. Then the state vector is modeled as a vector autoregressive

moving-average (VARMA), which can later be extended to include conditional information from other explanatory variables. Since $\tilde{\mathbf{x}}_i$ is a vector of only ones and zeros, it should also be that $0 \leq E[\tilde{\mathbf{x}}_i] \leq 1$. To directly impose this condition for any set of covariates, the researchers apply the logistic link function to express the VARMA model in terms of the log odds of the price change states with respect to a base state. Given the linear structure of the VARMA model, the base state can be chosen arbitrarily without a loss of generality. By doing so, they can then construct a $(k-1) \times 1$ vector of conditional probabilities, $\boldsymbol{\pi}_i$, where the conditional probability of the k^{th} state can be found by setting $\sum_{m=1}^k \pi_{im} = 1$. Defining a vector of the log probability ratios, they let

$$\mathbf{h}(\boldsymbol{\pi}_i) = \ln \left(\frac{\boldsymbol{\pi}_i}{1 - \boldsymbol{\nu}'\boldsymbol{\pi}_i} \right) = \mathbf{P}\mathbf{x}_i + \mathbf{c}, \quad (16)$$

where $\boldsymbol{\nu}$ is a conforming vector of ones, \mathbf{P} is an unspecified $(k-1) \times (k-1)$ time-invariant transition matrix, \mathbf{x}_i is the $(k-1) \times 1$ state vector, and \mathbf{c} is a $(k-1)$ dimensional vector of constants.

By generalizing Equation (15) to allow \mathbf{P} to consist of time-varying transition probabilities and by expanding the dependent information set, Engle and Russell obtain a model that is much richer and dynamic in structure. The so-called Autoregressive Conditional Multinomial (ACM) model of order (p, q, r) is then given by

$$\mathbf{h}(\boldsymbol{\pi}_i) = \sum_{j=1}^p \mathbf{A}_j(\mathbf{x}_{i-j} - \boldsymbol{\pi}_{i-j}) + \sum_{j=1}^q \mathbf{B}_j \mathbf{h}(\boldsymbol{\pi}_{i-j}) + \boldsymbol{\chi}\mathbf{v}_i \quad (17)$$

Where \mathbf{A}_j and \mathbf{B}_j denote the j^{th} $(k-1) \times (k-1)$ parameter matrices; $\mathbf{v}_i = [1 \ v_1 \ \dots \ v_r]'$ an $(r+1)$ -dimensional vector consisting of 1 in the first element to form a constant and the v_l for $l = 1, \dots, r$ are explanatory variables, and $\boldsymbol{\chi}$ is a $(k-1) \times (r+1)$ parameter matrix. In their paper, Engle and Russell specify the explanatory variables to be r -lags of trade duration, albeit they mention other possibilities such as trade volume and prevailing bid-ask spread. The terms $\{\mathbf{x}_i - \boldsymbol{\pi}_i\}$ form a martingale difference sequence describing the innovation associated with the i^{th} trade where \mathbf{A}_j determines its impact and \mathbf{B}_j can be interpreted as the rate of decay for past trade information. As we have seen before, the conditional probabilities, $\boldsymbol{\pi}_i$, can be obtained through logistic transformation.

4.4 Decomposition-ACM

Asset prices at the high frequency level live on a discrete grid and tend to exhibit strong temporal dependencies. As such, a robust model for price changes must be capable of capturing these characteristics and flexible enough to consider a range of explanatory trading variables. Thus far we have examined two models that are highly capable of describing the tick level price process, the Decomposition model and the ACM model. The Decomposition model is successful in that by parsing the price movement process, Z_i , specific trading phenomena such as increased price sensitivity to order flow, bid-ask bounce, and mean-reverting prices can be analyzed that are otherwise not immediately apparent in directly constructed high frequency price models such as the ACM. However, the model fails slightly in terms of a distributional misspecification for the size process, S_i and it is in this area that the ACM model of Engle and Russell succeeds. The VARMA structure with time-varying parameters and martingale difference sequence innovations in the ACM approach provides for a rich, flexible model that allows the price transition probabilities to be easily interpreted. In this section, we construct a new model for high-frequency price movements that features the robustness of the ACM while capturing the additional microstructure information obtained through decomposition by considering the price activity and direction series. Moreover, in addition to developing the model, which we term the Decomposition-ACM, some theoretical properties and estimation procedures are also established.

4.4.1 Model Specification

Recall from equation (1), we define the high frequency price of a financial asset at a specific instance in time to be the random sum $p(t) = p(0) + \sum_{i=0}^{N(t)} Z_i$, where t is a continuous clock. In this framework, the price process $p \equiv \{p(t) : t \geq 0\}$ can be thought of as a compound Poisson process. That is, a continuous-time stochastic process with jumps arriving randomly generated by a Poisson process $N \equiv \{N(t) : t \geq 0\}$ with rate parameter $\lambda > 0$. The jumps $Z_i \equiv \{z_i : z_i \in \mathbb{Z}, i \geq 1\}$ correspond to the change in the price of the asset induced by the i^{th} trade and possess their own

interesting probability distribution. In the literature it has been shown that the price process, $p(t)$, can be sufficiently characterized by considering the joint distribution of the price movements and the trade arrivals*. Consequently, our attention is directed toward formulating and estimating the joint conditional distribution of the Z_i given by equation (2) in terms of our Decomposition-ACM model (we discuss the trade arrival process in greater detail in section 5). Following the original proposal by Rydberg et al. (2003), we decompose the Z_i into a trivariate mixture model with price change activity, direction, and size (for reference, see section 4.1). Through this transformation we obtain the probability distribution of Z_i conditional on the σ -field \mathcal{F}_{i-1} , given as previously posed in equation (7), as $P(Z_i|\mathcal{F}_{i-1}) = P(A_i|\mathcal{F}_{i-1})P(D_i|A_i, \mathcal{F}_{i-1})P(S_i|D_i, A_i, \mathcal{F}_{i-1})$. Note that in addition to having access to the information provided in \mathcal{F}_{i-1} , S_i is both contemporaneously dependent on the direction and activity while D_i is contemporaneously dependent on activity. We re-emphasize the intuitive, natural ordering present in the decomposition as it is an essential feature of this approach. Subsequently, it follows from equations (4), (5), and (7) that the conditional distribution for price movements,

$$\begin{aligned}
P(Z_i = z_i|\mathcal{F}_{i-1}) &= \mathbb{1}_{1i}(1 - p_i) + \mathbb{1}_{2i}p_i\delta_iP(S_i = z_i|D_i = 1, A_i = 1, \mathcal{F}_{i-1}) \\
&\quad + \mathbb{1}_{3i}p_i(1 - \delta_i)P(S_i = -z_i|D_i = -1, A_i = 1, \mathcal{F}_{i-1}) \quad (18)
\end{aligned}$$

Notice that equation (17) is identical to (13) except that we have left the conditional distribution for the size of the price movements unspecified. As previously mentioned in equation (10), Tsay (2010) defines the price magnitude process to be geometrically distributed and enforces response asymmetry by bifurcating the parameters for up and downward price changes. Although not shown, his stipulation is necessary to preserve the role of the D_i in the model as it establishes a fundamental distinction between the particular direction that the asset's price moved. Empirically, there is strong evidence to suggest that prices, do in fact, react asymmetrically in the presence of new information (see Rydberg et al. (2003), Engle (2000), Bowsher (2006)). Statistically significant direction lags in our model, in addition to the observed convergence in theory and empirical results, would further

enhance our argument for decomposing the Z_i . Let

$$P(S_i|D_i, A_i = 1, \mathcal{F}_{i-1}) = \boldsymbol{\pi}_i^l = \begin{bmatrix} \pi_{1i}^l \\ \vdots \\ \pi_{mi}^l \end{bmatrix}, \text{ where } l = u, d \quad (19)$$

where π_{ki}^l denotes the conditional probability that the i^{th} trade induces the S_i to transition to the k^{th} state dependent on whether the magnitude of the price change was in the up or downward direction. Each of the $k = 1, \dots, m$ corresponds to the particular magnitude of price change measured in ticks associated with the i^{th} trade. As in equation (15), we propose estimating the conditional probabilities by means of their log odds, yielding

$$\mathbf{h}(\boldsymbol{\pi}_i^l) = \ln \left(\frac{\boldsymbol{\pi}_i^l}{1 - \boldsymbol{\nu}' \boldsymbol{\pi}_i^l} \right) = \mathbf{T}^l \mathbf{s}_i + \mathbf{c}^l \quad (20)$$

where $\boldsymbol{\pi}_i^l$ is now a $(m-1)$ -dimensional vector since the probability ratios are taken with respect to a base state and \mathbf{T}^l , \mathbf{s}_i , and \mathbf{c}^l play identical roles to their counterparts in (15). In our variation, we let the set of the magnitudes of tick changes, $M_{\Delta p_i} = \{0, 1, 2, 3\} \rightarrow \{1, 2, 3\}$. Note that our state space is limited to contain only magnitudes and not direction as this variable has already been taken into account previously in the decomposition. Without a loss of generality, the state indicating a change of zero is chosen to serve as the base state. We picked this state due to the natural ordering in our model, which differentiates between zero and non-zero price changes. Since we are constructing a model to estimate $P(S_i|D_i, A_i = 1, \mathcal{F}_{i-1})$, it makes sense to measure the likelihood of a non-zero change relative to no change considering this probability is only non-trivial when $S_i \neq 0$. Then \mathbf{s}_i assumes the j^{th} column of the identity matrix, \mathbf{I}_M , when the j^{th} state of $M_{\Delta p_i}$ occurs so that

$$\mathbf{s}_i = \begin{cases} [1 \ 0 \ 0]', & \text{if } |\Delta p_i| = 1 \\ [0 \ 1 \ 0]', & \text{if } |\Delta p_i| = 2 \\ [0 \ 0 \ 1]', & \text{if } |\Delta p_i| = 3 \end{cases}$$

Furthermore, since \mathbf{s}_i is distributed multinomially, the form of its conditional covariance matrix is easily inferred (see MacRae 1977) to be

$$COV_i \equiv Cov(\mathbf{s}_i|\mathcal{F}_{i-1}) = diag\{\boldsymbol{\pi}_i\} - \boldsymbol{\pi}_i \boldsymbol{\pi}_i' = \begin{pmatrix} \pi_{1i}(1 - \pi_{1i}) & -\pi_{1i}\pi_{2i} & -\pi_{1i}\pi_{3i} \\ -\pi_{2i}\pi_{1i} & \pi_{2i}(1 - \pi_{2i}) & -\pi_{2i}\pi_{3i} \\ -\pi_{3i}\pi_{1i} & -\pi_{3i}\pi_{2i} & \pi_{3i}(1 - \pi_{3i}) \end{pmatrix} \quad (21)$$

Applying the ACM structure of (16) originally developed by Engle et al. (2005), $\mathbf{h}(\boldsymbol{\pi}_i^l)$ takes the form

$$\mathbf{h}(\boldsymbol{\pi}_i^l) = \sum_{j=1}^p \mathbf{B}_j^l (\mathbf{s}_{i-j} - \boldsymbol{\pi}_{i-j}^l) + \sum_{j=1}^q \mathbf{C}_j^l \mathbf{h}(\boldsymbol{\pi}_{i-j}^l) + \sum_{j=1}^r \boldsymbol{\beta}_j A_{i-j+1} + \sum_{j=1}^r \boldsymbol{\gamma}_j D_{i-j+1} + \mathbf{c}^l \quad (22)$$

where the A_i and D_i represent the contemporaneous and $(r-1)$ -lag values of activity and direction that we are familiar with from the Decomposition model of Rydberg et al. (2003) and $\boldsymbol{\beta}_j, \boldsymbol{\gamma}_j$ are $(k-1) \times r$ parameter matrices. This formulation is marked by a number of advantageous in comparison to the original ACM described in (16). By truncating the \mathbf{s}_i through the decomposition, we reduce the number of conditional probabilities to estimate at each trade by half, helping the model maintain parsimony. Furthermore, the inclusion of the activity variable should aid the computational efficiency of the model since the probability of an inactive trade is now found directly instead of having to find the residual probability from all the other possible states. Reinforcing the importance of this stipulation empirically, inactive trades are common, particularly during periods of low volatility. Now by applying the logistic transformation to (20) we arrive at our desired conditional distribution for the size of price change induced by the i^{th} trade and have that

$$P(S_i | D_i, A_i = 1, \mathcal{F}_{i-1}) = \boldsymbol{\pi}_i^l = \frac{\begin{bmatrix} \pi_{1i}^l \\ \vdots \\ \pi_{mi}^l \end{bmatrix}}{1 + \boldsymbol{\nu}' e^{\mathbf{h}_i^l}} \quad (23)$$

Combining this result with that of equation (17), we obtain the conditional distribution of price movements,

$$P(Z_i = z_i | \mathcal{F}_{i-1}) = \mathbb{1}_{1i}(1 - p_i) + \mathbb{1}_{2i} p_i \delta_i \mathbf{s}_i' \boldsymbol{\pi}_i^u + \mathbb{1}_{3i} p_i (1 - \delta_i) \mathbf{s}_i' \boldsymbol{\pi}_i^d \quad (24)$$

The computational benefits of this model become more apparent in this formulation. It can easily be seen that our indirect approach allows us to only have to focus on the relevant component of (22) for each trade, rather than the direct approach, which requires to tackle the whole problem at once.

4.4.2 Estimation

Given the assumptions of our model and equation (22), we can construct the log likelihood function as a product of the conditional densities to be,

$$L(\Theta) = \sum_{i=1}^N \ln[P(Z_i = z_i | \mathcal{F}_{i-1})], \quad (25)$$

where Θ represents the set of parameters to be estimated. For the case that $Z_i = 0 (A_i = 0)$, the current model is identical to the one estimated in section 4.2.1, which was conducted via maximum likelihood as the regularity conditions for the logistic distribution are well understood. The estimated model only deviates from that in section 4.2.1 when $Z_i = z_i \neq 0$ since the conditional distribution of the size of the price change process is redefined using the Decomposition-ACM. As noted by Engle et al. (2005), the ACM(p,q) model is analogous in structure to the more familiar GARCH(p,q) process. Consequently, equation (21) considered independently has a log likelihood function whose partial derivatives assume the recursive form present in GARCH models originally demonstrated by Bollerslev (1986), which are shown to produce consistent, efficient, asymptotically normal maximum likelihood parameter estimates. Therefore, we assume that when the mixture model is considered in aggregate, the regularity conditions will be preserved and we will obtain consistent, efficient maximum likelihood estimates for our parameters. To conduct this estimation procedure, suggest implementing the Berndt, Hall, Hall, and Hausman (BHHH) (1974) numerical optimization algorithm. Although rather computationally inefficient for optimization on large datasets, as in this paper, Bollerslev (1986) notes that the recursive structure of the log likelihood derivatives conveniently fit the BHHH procedure.

5 Trade Arrivals

Among the myriad peculiarities that differentiates the study of high frequency finance from its lower frequency counterpart, perhaps the most salient is the irregular spacing of data in time. Usually, sequences of asset prices are considered over aggregated fixed-intervals to facilitate analysis, so the

issue of handling random transaction data becomes inconsequential. However, this turns out to be a costly simplification at the trade-by-trade level. It has been well discovered that the timing of trading events, particularly the arrivals of trades and the frequency in which they occur, possess indispensable information for market microstructure analysis and intraday volatility forecasting (see Bauwens and Hautsch (2007), Engle and Russell (1998)). Hence, it is integral that we construct a model that can accurately depict and preserve the features of these trade arrivals to fully characterize the price process, $p(t)$. As discussed in the literature review, the typical approach for this task involves the implementation of a so-called financial point process. Essentially, these are continuous time point processes with a memory of past trading events. In the literature, models have been considered from two vantage points of the trading process, duration and intensity. Intensity based models are attractive in that they naturally suit continuous-time modeling in the univariate and multivariate framework. A possible extension to this paper could be to construct a high frequency mean-variance efficient portfolio. If we elected to implement a duration model, we would encounter the well-known dilemma in finance of matching asynchronous durations, which would vastly inhibit our ability to consistently estimate the portfolio's intraday covariance matrix of volatilities and cross-volatilities. In this paper, we develop the financial point process from the more flexible intensity perspective (Russell 1999) by modeling it as a so-called Cox process.

5.1 Trading Intensity

While the definition of trading intensity will become more clear as we develop the mathematics, initially it can be thought of as the instantaneous probability that an asset is traded (a trade arrival). Fundamentally, the analysis of trade arrivals is rooted in point process theory and is the starting point in our development of the model. Let $\{t_i\}_{i=1}^n$ be a monotone increasing random sequence of event times and let $N \equiv \{N(t) : t \geq 0\}$ be a càdlàg counting function. We say that $N(t)$

is a non-homogenous Poisson process (NHPP) with respect to the mean measure, $\Lambda(t)$, iff

- (1) $P(N(0) = 0 | \mathcal{F}_0) = 1$
- (2) $\forall t, s \geq 0, 0 \leq u \leq t, N(t+s) - N(t)$ is independent from $N(u)$
- (3) $\forall t, s \geq 0, P(N(t+s) - N(t) = 1 | \mathcal{F}_t) = \lambda(t)s + o(s)$
- (4) $E[N(t) | \mathcal{F}_t] = \Lambda(t) = \int_0^t \lambda(s) ds < \infty$
- (5) Increments, $\tau_i = t_i - t_{i-1}$, are independent, but are not stationary

where,

$$\lambda(t | \mathcal{F}_t) = \lim_{h \rightarrow 0^+} \frac{1}{h} E[N(t+h) - N(t) | \mathcal{F}_t] \quad (26)$$

is called the \mathcal{F}_t -conditional intensity of $N(t)$. One should also note that the \mathcal{F}_t -conditional process $N(t)$ is a submartingale, that is $E[N(t) | \mathcal{F}_s] \geq N(s), s < t$, with compensator $\Lambda(t)$. A Cox process, or doubly-stochastic Poisson process, is a generalization of the NHPP where the intensity function $\lambda(t)$ is defined to be its own random process in such a way that $N(t) | \lambda(t) \sim NHPP(\lambda(t))$ and $\lambda(t)$ becomes an \mathcal{F}_s -predictable function. In the literature, the stochastic $\lambda(t)$ is represented in a variety of forms such as an autoregressive process as in Hamilton, Jordà (2002) or Russell (1999), or as an Ornstein-Uhlenbeck process in Rydberg, Shephard (1998). Autoregressive intensity models are successful in that they are able to capture a variety of features in the data such as transaction clustering (Hamilton et al. 2002). Rather than specifying a parametric intensity model, we elect to take a nonparametric approach in constructing our estimate for the density of the trading process. Although we lose the ability to specific dependencies in the underlying process, such as those inducing trade clustering, are model is less impacted by bias associated with the distributional and structural assumptions made about $\lambda(t)$. Trading intensity tends to exhibit both short term and long term dependencies, making it potentially difficult to construct a parsimonious parametric model without a comprehensive prior understanding of the process. While in a biophysical setting, Zhang and Kou (2010) provide a framework for nonparametric estimation and inference of Cox processes via kernel density estimation that can simply be adapted to fit our context of trade

arrivals. Furthermore, Zhang et al. (2010) estimate the process's autocorrelation function (ACF), which will become instrumental in later specifying a more descriptive parametric model.

5.1.1 Nonparametric Estimation

Let $t_1 < t_2 < \dots < t_n$, $t \in [0, T]$ denote a random sequence of increasing arrival times from a Cox process with stochastic intensity λ . Then the Rosenblatt-Parzen estimator for the intensity (density) λ estimated at $s \in \mathbb{R}$ is,

$$\hat{\lambda}_h(s) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{s - t_j}{h}\right) \quad (27)$$

where $h > 0$ is the smoothing bandwidth and K is a symmetric kernel satisfying

$$\int_{\mathbb{R}} K(s) ds = 1$$

A main factor in determining the performance of this estimator is the choice of bandwidth, h . Zhang and Kou (2010) propose optimization this selection by minimization of the mean integrated squared error (MISE) and a relatively simple regression plug-in method. Assuming that the true realization $\lambda(t)$ is ergodic, an estimate for the process's autocorrelation function can be easily constructed once obtaining $\hat{\lambda}_h(t)$.

6 Conclusion

In this paper, we discuss the characteristics of security prices at the trade-by-trade level and their dissimilarity to prices observed over longer fixed intervals. Moreover, we describe the frequency in which transaction data is accumulated throughout a trading day and its implication for asset prices. The relationship between trading intensity and financial returns is prominent topic among market microstructure theorists as in the works of Easley et al. (1992) and Diamond et al. (1987), which serve as a motivation for the analysis of high frequency data. We propose an unique econometric methodology capable of preserving the irregularity of transaction data by defining the price process as a compound Poisson process. Following the approach of Rydberg et al. (2003), we decompose the price movement process into a naturally ordered trivariate mixture model of activity, direction,

and magnitude. Activity, which indicates a trade induced price change, and direction are modeled as an autoregressive logistic process. We extend the Rydberg and Shephard model by describing the magnitude process as a more dynamic finite-state VARMA process. The VARMA approach was originally proposed by Engle and Russell (2005), terming their specific model the autoregressive conditional multinomial (ACM). However, they choose to model high frequency price movements directly as opposed to Rydberg et al. (2003), who do so indirectly. Approaching the problem indirectly through decomposition is a more informative framework as it is able to uncover relationships which are not apparent in direct modeling. This lead us to develop the so-called decomposition-ACM model for high frequency price movements. Additionally, we allude to some of its properties and potential estimation procedures for the model. Furthermore, we describe the trade arrivals as a Cox process and propose estimating its random intensity through nonparametric kernel methods. For future research, the asymptotic properties of the decomposition-ACM can be explored and our methodological framework can be implemented to study the factors contributing to a security's price formation process. Interesting extensions could also be generalizing our model to the multivariate case in order to engineer an efficient high frequency portfolio and to consider other areas of application for our model such as in constructing the stock price lattice used in multinomial option pricing.

7 Acknowledgments

I could not be more grateful for the time of my thesis advisor, Carlos Martins-Filho; for without his patience and mentorship this paper would not be possible. I would also like to thank Chris Leach for granting me access to the NYSE TAQ data. Lastly, I want to thank Daniel Brown for being more than accommodating in assisting me with the final stages of the thesis on such short notice.

References

- [1] Ait-Sahalia, Y., R.F. Engle, L.P. Hansen, and J.R. Russell (2010): "Analysis of High-Frequency Data." *Handbook of Financial Econometrics: Tools and Techniques*. San Diego: North-Holland, 383-426.
- [2] Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008): "Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise." *Econometrica* 76.
- [3] Bauwens, L. and N. Hautsch (2007): "Modeling Financial High Frequency Data Using Point Processes." *Handbook of Financial Time Series*.
- [4] Berndt, E., B. Hall, R. Hall, and J. Hausman (1974): "Estimation and Inference in Nonlinear Structure Models." *Annals of Economic and Social Measurement* 3, 654-665.
- [5] Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31, 307-327.
- [6] Bowsher, C. G. (2006): "Modeling Security Markets in Continuous Time: Intensity Based Multivariate Point Process Models." *Journal of Financial Econometrics*.
- [7] Brownlees, C. T., and G. M. Gallo (2006): "Financial Econometric Analysis of Ultra-High Frequency: Data Handling Concerns." *Computational Statistics and Data Analysis* 51, 2232-2245.
- [8] Cox, D. R. (1972): "Regression Models and Life Tables." *Journal of the Royal Statistical Society Series B*, 34,187-220.
- [9] Diamond, D.W., and R.E. Verrecchia (1987): "Constraints on Short-Selling and Asset Price Adjustments to Private Information." *Journal of Financial Economics* 18, 277-311.

- [10] Easley, D., and M. O'Hara (1992): "Time and the Process of Security Price Adjustment." *The Journal of Finance* 19, 69-90.
- [11] Engle, R. F. (2000): "The Econometrics of Ultra-High-Frequency Data." *Econometrica*, 1-22.
- [12] Engle, R. F., and J.R. Russell (2005): "A Discrete-State Continuous Time Model of Financial Transaction Prices and Times: The Autoregressive Conditional Multinomial Autoregressive Conditional Duration Model." *Journal of Business and Economic Statistics* 23, 166-180.
- [13] Engle, R. F., and J. R. Russell (1998): "Autoregressive Conditional Duration: A New Model for Irregularly spaced Transaction Data." *Econometrica* 66, 1127-1163.
- [14] Gouriéroux, C. and J. Jasiak (2001): *Financial Econometrics: Problems, Models, and Methods*. New Jersey: Princeton University Press.
- [15] Hamilton, J. D., and O. Jorda (2002): "A Model of the Federal Funds Rate Target." *Journal of Political Economy* 110, 1135-1167.
- [16] Hasbrouck, J. (2007): *Empirical Market Micro-Structure: The Institutions, Economics, and Econometrics of Securities Trading*. Oxford University Press.
- [17] Hasbrouck, J. (1991): "Measuring the Information Content of Stock Trades." *The Journal of Finance* 66, 179-207.
- [18] MacRae, E. (1997): "Estimation of Time-Varying Markov Processes with Aggregate Data." *Econometrica* 45, 183-198.
- [19] McFadden, D. and W. Newey (1986): "Large Sample Estimation and Hypothesis Testing." *Handbook of Econometrics* 4, 2111-2245.
- [20] Poon, S., and C.W.R. Granger (2003): "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 66, 478-539.

- [21] Russell, J. R. (1999): "Econometric Modeling of Multivariate Irregularly-Spaced High-Frequency Data." Working Paper, University of Chicago.
- [22] Rydberg, T. H., and N. Shephard (2003): "Dynamics of Trade-by-Trade Price Movements: Decomposition and Models." *Journal of Financial Econometrics* 1, 2-25.
- [23] Sheather, S. J. (2004): "Density Estimation." *Statistical Science* 4, 588-597.
- [24] Shephard, N. (1995): "Generalized Linear Autoregressions." Unpublished Manuscript, Nuffield College, Oxford.
- [25] Tsay, R.S. (2010): *Analysis of Financial Time Series*. New Jersey: John Wiley & Sons.
- [26] Zhang, T. and S. C. Kou (2010): "Nonparametric Inference of Doubly Stochastic Poisson Process Data via the Kernel Method." *The Annals of Applied Statistics* 4, 1913-1941.
- [27] Zhou, B. (1996): "High Frequency Data and Volatility in Foreign Exchange Markets." *Journal of Business and Economic Statistics* 14, 45-52.