Autoregressive Conditional Duration:
A New Model for Irregularly Spaced Transaction Data

By Robert F. Engle
University of California, San Diego
and
Jeffrey R. Russell
University of Chicago, Graduate School of Business

First Draft: December 1994
Current Draft: July 1997

Abstract

This paper proposes a new statistical model for the analysis of data which arrives at irregular intervals. The model treats the time between events as a stochastic process and proposes a new class of point processes with dependent arrival rates.   The conditional intensity is developed and compared with other self-exciting processes. Because the model focuses on the expected duration between events, it is called the Autoregressive Conditional Duration (ACD) model. Asymptotic properties of the Quasi Maximum Likelihood estimator are developed as a corollary to ARCH model results. Strong evidence is provided for duration clustering for the financial transactions data analyzed; both deterministic time-of-day effects and stochastic effects are important. The model is applied to the arrival times of trades and therefore is a model of transaction volume, and also to the arrival of other events such as price changes. Models for the volatility of prices are estimated with price based durations, and examined from a market microstructure point of view.

---

1. Introduction

With the rapid development in computing power and storage capacity, data is being collected and analyzed at ever higher frequencies. For many types of data, the ultimate in high frequency data collection has been reached and every transaction is recorded. This limit has been reached for financial market transactions which are the focus of this paper, as well as telephone calls, credit card purchases, or the sale of any good using scanning devices. Since the quantity purchased in a period of time is often the key economic variable to be modeled or forecast, it is natural to study the timing of transactions. Financial market microstructure theories are typically tested on a transaction by transaction basis so again the timing of these transactions can be central to understanding the economics.

Transaction data inherently arrive in irregular time intervals, while standard econometric techniques are based on fixed time interval analysis. There is a natural inclination for the econometrician to aggregate transactions data to some fixed time interval. For purchases of consumer durables by an individual, a natural interval might be months or even years. On the other extreme, frequently traded stocks will have transactions every few seconds, hence a much shorter interval is appropriate. If a short time interval is chosen, there will be many intervals with no new information and heteroskedasticity of a particular form will be introduced into the data. On the other hand, if a long interval is chosen, the micro structure features of the data will be lost. In particular, multiple transactions will be averaged and the characteristics and timing relations of individual transactions will be lost, mitigating the advantages of moving to transaction data in the first place.

The problem becomes more complicated when one realizes that the rate of arrival of transactions type data may vary over the course of the day, week or year making the choice of an "optimal" interval more difficult. For stocks, activity is higher near the open and the close than in the middle of the day. For currency markets, there are clear periods of high and low activity as markets around the world open and close. Even more intriguing is the case of transactions which are generally infrequent but which may

suddenly exhibit very high activity. This may be due to some observable event such as a news release or to an unobservable event which may best be thought of as a stochastic process. In these cases the choice of a fixed interval for data analysis is very perilous as it may leave the investigator with many uninformative points, or disguise the periods of most interest.

This paper will propose an alternative to fixed interval analysis. The arrival times are treated as random variables which follow a point process. Associated with each arrival time are random variables called *marks* such as volume, bid ask spread, or price. A new model for dependent point processes is formulated. The conditional intensity function is parameterized in terms of past events in a way that seems particularly well suited for the transactions process. The most fundamental application of the model is to measure and forecast the intensity of transaction arrivals which is essentially the instantaneous quantity of transactions. The basic formulation of the model parameterizes the conditional intensity as a function of the time between past events, and numerous natural extensions include other effects such as characteristics associated with past transactions, or any other outside influence. The dependence of the conditional intensity on past durations suggests that the model be called the Autoregressive Conditional Duration (ACD) model.

Sometimes the investigator may not be interested in modeling the time between transactions but rather in studying the marks associated with the arrival times. A method is proposed for modeling the rate of change of other variables by selectively thinning the point process. A model for the intensity of the price changes is developed for the transactions data analyzed. In this context, hypotheses from market microstructure theories can be examined.

Associated with the intensity is the conditional expectation of the waiting time until the next event. The model therefore has an interesting interpretation in the context of time deformation models because the model is formulated in transaction time but models the frequency and distribution of calendar time between events. See for example, Clark (1973) and Tauchen and Pitts (1983), and more recently Stock (1988), Muller et al. (1990) and Ghysels and Jasiak (1994). The ACD formulation can use but does not require

auxiliary data or assumptions on the causes of time flow; it is simply a time series model of time.

The model is applied to transactions data from financial markets. Recent research by Kyle (1985), Admati and Pfleiderer (1988) and Easly and O'Hara (1992) has suggested that the frequency of transactions should carry information about the state of the market. In particular the models suggest clustering of transactions. Clustering of transactions is exactly what is found by examining IBM transactions data. Even after time-of-day effects are removed, large autocorrelations exist in the time intervals between trades. The ACD model also provides a framework to test if the intensity function is influenced by observed market variables. In particular, several models are estimated in an attempt to identify the source of transaction clustering.

The following section develops the statistical underpinning for dependent point processes and section 3 introduces the Autoregressive Conditional Duration model. Section 4 discusses asymptotic properties of the ACD model and section 5 discusses extensions of the model. Section 6 gives empirical results for IBM transaction data which in Section 7 are related to the economics of market microstructure. Section 8 discusses an extension of the model to model the price process and test some hypothesis about market microstructure for the IBM data. Finally, section 9 concludes.

2. The Conditional Intensity Process

This section will give a brief description of some of the relevant point process models for intertemporally correlated events.   After this discussion, the particular parameterization of the ACD model will be developed.

Consider a stochastic process that is simply a sequence of times $\{t_0, t_1, ..., t_n, ...\}$ with $t_0 < t_1 < ... < t_n ...$.   Associated with the arrival times is the counting function $N(t)$ which is the number of events that have occurred by time t. Clearly it is a step function which is continuous from the left with limits from the right. If there are characteristics associated with the arrival times, such as a price or volume, the process is called a "*marked point process*".

Two general characterizations of a point process can be introduced here following Snyder and Miller(1991). A point process on $[t_0, \infty)$ is said to "*evolve without aftereffects*" if for any $t > t_0$, the realization of points during $[t, \infty)$ does not depend in any way on the sequence of points during the interval $[t_0, t)$. A counting process is said to be "*conditionally orderly*" at time $t \geq t_0$ if for a sufficiently short interval of time and conditional on any event P defined by the realization of the process on $[t_0, t)$, the probability of two or more events occurring is infinitessimal relative to the probability of one event.

We focus on point processes which evolve *with* aftereffects and which are conditionally orderly. A complete description of such processes is naturally formulated in terms of the intensity function conditional on all available past information which must minimally include the arrival times and the count. This *conditional intensity process* is therefore defined[1] by

$$(1) \qquad \lambda(t | N(t), t_1, \ldots, t_{N(t)}) = \lim_{\Delta t \to 0} \frac{P(N(t + \Delta t) > N(t) | N(t), t_1, \ldots, t_{N(t)})}{\Delta t}$$

In many applications this will equivalently be called the hazard function, particularly in a context where there may be many individuals rather than a point process under study. As discussed in Lancaster (1990) or Snyder and Miller(1991) for example, the conditional intensity, the conditional density of the durations or *"waiting times"* between events, and the conditional survivor function each are complete descriptions of a conditionally orderly stochastic process. Letting p_i be a family of conditional probability density functions for arrival time t_i the log likelihood can be expressed in terms of the conditional densities or intensities as:

$$(2) \qquad \begin{aligned} L &= \sum_{i=1}^{N(T)} \log p_i(t_i | t_0, \ldots, t_{i-1}) \\ L &= \sum_{i=1}^{N(T)} \log \lambda(t_i | i - 1, t_0, \ldots, t_{i-1}) - \int_{t_0}^{T} \lambda(u | N(u), t_0, \ldots, t_{N(u)}) du \end{aligned}$$

---

[1] Throughout it will be assumed that when N(t)=0, there are no further arguments to the function.

Equation (1) is a general statement of the intensity function of a "self-exciting" point process which is a process where the past evolution impacts the probability structure of future events. It was originally proposed by Hawkes (1971a,b 1972) and by Rubin(1972). These are sometimes called Hawkes self-exciting processes. The success in using such processes depends upon the parameterization of the conditional intensity.

The simplest point process in this class is the *Poisson process* for which λ is a constant parameter. A more flexible process is the *Inhomogeneous Poisson process* for which the intensity varies only with t itself so that the arrival rate is assumed to be a deterministic function of time. In neither case however, do past events influence the future arrival rates; they evolve without aftereffects.

When the intensity depends on the number of events but not the timing of these events, then the process is a *pure Birth process*. Rubin(1972) introduced limited memory self-exciting processes. A process is called an "*m-memory self-exciting counting process*" if only the m most recent arrival times are present in the conditional intensity. In this notation, a zero memory self-exciting process is a *Markov birth process*, and a homogeneous 1-memory process is a *renewal process*.

With longer memory there are many suggestions on how to parameterize the conditional intensity. We will briefly describe two existing classes of point process models that are characterized by (1). The first class of models are formulated in calendar time. Linear representations of this class of models can be expressed as:

$$(3) \qquad \lambda\left(t \middle| N(t), t_1, \ldots, t_{N(t)}\right) = \omega + \sum_{i=1}^{N(t)} \pi(t - t_i)$$

where each past arrival time t_i contributes $\pi(t-t_i)$ to the intensity at time t. π is called an infectivity measure as motivated by epidemiology, as well as population dynamics and earthquake prediction[2]. These types of specifications were initially proposed by Hawkes (1971a,b 1972) and are used in Ogata and Akaike (1982) and Vere-Jones and Ozaki (1982) with Laguerre polynomials defined on t>0 for the infectivity measure. These calendar time models are inappropriate in the study of transactions type data since they

---

[2] See Ogata and Katsura (1986) for a extensive review.

imply that the marginal effect of an event that occurred, say, 10 minutes in the past is independent of the intervening history; there may have been 0 events or 100 events in between.

The second class of conditional intensity parameterizations focus on the intervals between events and are formulated in event time. In these, the conditional intensity could be parameterized in terms of

$$(4) \qquad \lambda\left(t \mid N(t), t_1, \ldots, t_{N(t)}\right) = \omega + \sum_{i=1}^{N(t)} \pi_i (t_{N(t)+1-i} - t_{N(t)-i})$$

so that the impact of a duration between successive events depends upon the number of intervening events. Such models appear to have first been studied by Wold (1948) and later by Cox (1955). Wold proposed a model for correlated intervals using an autoregressive structure similar to standard ARMA techniques. The model was subsequently reformulated by Gaver and Lewis(1980), Lawrence and Lewis(1980) and Jacobs and Lewis(1977) as "*exponential autoregressive-moving average EARMA(p,q) models.*" These models assume that the durations are conditionally exponentially distributed with a mean which follows an ARMA process. The formulation of an additive error process which has this property is complex and the resulting maximum likelihood procedures are virtually unworkable, at least in general settings.

In some cases, the conditional intensity can be derived from more fundamental assumptions. For example, the *Cox(1955) or Doubly Stochastic* model typically assumes that there is a latent independent process which governs the arrival rate[3]. Suppose this is a counting process M(t) which might be called "information" for financial applications. Thus the intensity is conditional on M(t) as well as t itself. Snyder and Miller(1991) in Theorem 7.2.2 prove that such a process is itself a self-exciting process with a form given by (1), which is simply the expectation of the intensity over M conditional on the past history of N. Hence, the class of self-exciting processes includes the Cox process although it is not generally easy to determine the form of the intensity process from the conditional expectation.

---

[3] See Grandell(1976) for a good discussion of the Doubly Stochastic models.

Cox (1955) introduced a model for duration analysis with covariates and later generalized the model to the popular proportional hazards framework. The conditional intensity can be written as

$$(5) \qquad \lambda\left(t \middle| z_{N(t)},\ldots,z_1\right) = \lambda(t) \exp\left\{\beta' z_{N(t)}\right\}$$

where z_i is a vector of explanatory variables associated with the arrival time i. One suggestion mentioned in Cox(1972b) was to include lagged durations as an explanatory variable. Duration models were introduced into econometrics by Lancaster(1979), and given a dynamic focus in Heckman and Borjas (1980) and Heckman (1981) among others, to examine the impact of past unemployment on current spells. In these models, the data are typically short time series on many individuals, so that the question of whether this truly reflects state dependence or merely unmeasured heterogeneity becomes very important.

In this paper, a new family of self-exciting processes will be introduced. They will have conditional intensities different from all of these representations. The process will be introduced in terms of its conditional densities.

3. The ACD model

The ACD model is most conveniently specified in terms of the conditional density of the durations. Letting $x_i = t_i - t_{i-1}$ be the interval between two arrival times which will be called the duration, the density of x_i conditional on past x's will be specified directly. Let ψ_i be the expectation of the ith duration which is given by

$$(6) \qquad E(x_i | x_{i-1},\ldots,x_1) = \psi_i(x_i | x_{i-1},\ldots,x_1;\theta) \equiv \psi_i.$$

Let the ACD class of models consist of parameterizations of (6) and the assumption that

$$(7) \qquad x_i = \psi_i \varepsilon_i$$

where

(8) $\{\varepsilon_i\} \sim i.i.d.$ with density $p(\varepsilon;\phi)$, and θ and ϕ are variation free.[4]

It is natural to call the model Autoregressive Conditional Duration or ACD since the conditional expectation of the duration will depend upon past durations.

From expression (6) - (8) it is apparent that we now have a host of potential specifications for the ACD model each defined by different specifications for the expected durations and for the distribution of ε. To derive a general expression for the conditional intensity, let p_0 be the density function of ε and let S_0 be the associated survival function. Then define

(9) $$\lambda_0(t) = \frac{p_0(t)}{S_0(t)}$$

as the *baseline hazard* using the name popularized by the proportional hazard literature. Then the conditional intensity of an ACD model can be expressed in general as

(10) $$\lambda(t|N(t),t_1,\ldots,t_{N(t)}) = \lambda_0\left(\frac{t-t_{N(t)}}{\psi_{N(t)+1}}\right)\frac{1}{\psi_{N(t)+1}}$$

so that the past history influences the conditional intensity by both a multiplicative effect and a shift in the baseline hazard. This model is called an *accelerated failure time* model since the past information influences the rate at which time passes. Such models are natural in some medical examples where patients with particular characteristics move through a disease more rapidly than others. These models are also natural in finance. A large literature on *time deformation* suggests that sometimes time flows very rapidly in financial markets while in other periods it moves slowly. In this case, the rate of time flow depends on the past event arrival times through the function ψ.

The simplest version of the ACD assumes that the durations are conditionally exponential so that the baseline hazard is simply one and the conditional intensity is

(11) $$\lambda(t|x_{N(t)},\ldots,x_1) = \psi_{N(t)+1}^{-1}$$

---

[4] If $\theta \in \Theta$ and $\phi \in \Phi$, then $(\theta,\phi) \in \Theta \otimes \Phi$.

An m-memory conditional intensity would imply that only the most recent m durations influenced the conditional duration suggesting a possible specification:

$$(12) \qquad \psi_i = \omega + \sum_{j=0}^{m} \alpha_j x_{i-j}$$

A more general model without the limited memory characteristic is

$$(13) \qquad \psi_i = \omega + \sum_{j=0}^{m} \alpha_j x_{i-j} + \sum_{j=0}^{q} \beta_j \psi_{i-j}$$

which will be called an ACD(m,q) where the m and q refer to the orders of the lags.

This model is convenient because it allows various moments to be calculated by expectation regardless of the form of the baseline hazard. For example, the conditional mean of x_i is ψ_i, the conditional duration, but the unconditional mean is

$$(14) \qquad E(x_i) = \mu = \frac{\omega}{\left(1 - \sum(\alpha_j + \beta_j)\right)}$$

This is most easily seen by taking expectations of both sides of equation (13) although a dynamic analysis in the proof of lemma 1 of Engle and Russell(1995b) establishes that this exists only when all the roots of the associated difference equation lie outside the unit circle.

The simplest and often very successful member of this family is the EACD(1,1) where E signifies the exponential distribution for the errors.

$$(15) \qquad \psi_i = \omega + \alpha x_{i-1} + \beta \psi_{i-1} \text{ for } \alpha, \beta \geq 0, \ \omega > 0, \ \forall i, \ i = 1...N$$

In this model, the conditional variance of x is ψ_i^2 but the unconditional variance is given by :

$$(16) \qquad \sigma^2 = \mu^2 \left(\frac{1 - \beta^2 - 2\alpha\beta}{1 - \beta^2 - 2\alpha\beta - 2\alpha^2} \right)$$

Thus whenever α>0 the unconditional standard deviation will exceed the mean exhibiting "excess dispersion" as so often noticed in duration data sets.

The model in (13) can also be formulated as an ARMA(m,q) model for durations. Letting $\eta_i \equiv x_i - \psi_i$ which is a Martingale difference sequence by construction, the duration process can be expressed as

$$(17) \quad x_i = \omega + \sum_{j=0}^{\max(m,q)} (\alpha_j + \beta_j)x_{i-j} - \sum_{j=0}^{q} \beta_j \eta_{i-j} + \eta_i$$

which is an ARMA(m,q) process with highly non-Gaussian innovations.

Forecasts of waiting times can be computed directly from this representation using the conventional ARMA analytics. Thus it is simple to compute analytically the expected waiting time until the $i+k^{th}$ transaction occurs. If all roots of the associated polynomial are less than unity, then the duration process will be mean reverting and the impact of a given duration on future expected durations will die out exponentially. Since the transactions to be analyzed occur within seconds of each other, the persistence of shocks will be very limited in calendar time unless the roots are very close to unity.

The formulation in (17) is very close to the EARMA models mentioned above since the conditional durations are exponential and follow an ARMA process. The simplicity arises from using a multiplicative error in the ACD.

4. Asymptotic properties of the Exponential ACD Model

Readers who are familiar with the ARCH class of models will immediately recognize the relationship to models of conditional variance. The ACD(1,1) is analogous to the GARCH(1,1) and will have many of the same properties. Just as the GARCH(1,1) is often a good starting point, the ACD(1,1) seems like a natural starting point. However, just as there are many alternative volatility models, there are many interesting possibilities here. For recent surveys on ARCH models and lists of different classes, see Bollerslev, Engle and Nelson(1994), Bollerslev Chou and Kroner(1992) and Bera and Higgins (1992). The ARCH model was originally introduced by Engle(1982).

The connection with GARCH models is however even deeper than this analogy suggests. The theorems which establish QMLE properties of GARCH(1,1) even in the presence of unit roots can be carried over to the EACD(1,1) models as in the following corollary to Theorems 2 and 3 of Lumsdaine(1996) or Theorems 1 and 3 in Lee and Hansen(1994).

Corollary to Lee and Hansen(1994)

If

1) $\quad E_{i-1}(x_i) \equiv \psi_{0,i} = \omega_0 + \alpha_0 x_{i-1} + \beta_0 \psi_{0,i-1},$

$\quad\quad \varepsilon_i \equiv x_i / \psi_{0,i}$ is i) strictly stationary and ergodic

2)
$\quad\quad\quad\quad\quad\quad\quad$ ii) nondegenerate

$\quad\quad\quad\quad\quad\quad\quad$ iii) has bounded conditional second moments

$\quad\quad\quad\quad\quad\quad\quad$ iv) $\sup_i E[\ln(\beta_0 + \alpha_0 \varepsilon_i)|F_{i-1}] < 0 \quad a.s.$

3) $\quad \theta_0 \equiv (\omega_0, \alpha_0, \beta_0)$ is in the interior of Θ

4) $\quad L(\theta) = -\sum_{i=1}^{N(T)} \left(\log(\psi_i) + \frac{x_i}{\psi_i} \right)$ where $\begin{array}{l} \psi_i = \omega + \alpha x_i + \beta \psi_{i-1} \text{ for } i > 1 \\ \psi_i = \omega/(1-\beta) \text{ for } i = 1 \end{array}$

Then:

a) the maximizer of L will be consistent and asymptotically normal with a covariance matrix given by the familiar robust standard errors from Lee-Hansen, and

b) the model can be estimated with ARCH software by taking $\sqrt{x_i}$ as the dependent variable and setting the mean to zero.

Proof:

Let $y_i = d_i \sqrt{x_i}$ where d_i is independent of x and i.i.d. equal to 1 with probability .5 and -1 with probability .5.   Then the expected value of y_i is zero and its   variance is $\psi_{0,i}$. Substituting into condition 4 gives exactly the Gaussian quasi-loglikelihood used by Lee and Hansen. Conditions i) and ii) follow because the square root of a strictly stationary, ergodic and non-degenerate positive random variable will also be strictly stationary, ergodic and non-degenerate. Conditions iii) and iv) are equivalent to the conditions in Lee and Hansen. Thus a) follows from Hansen and Lee. Result b) follows from noticing that d_i disappears from the likelihood. QED.

If $\alpha + \beta$ is less than or equal to one, then condition (iv) of assumption 2 is automatically satisfied, but it is also satisfied for some unit root and explosive cases. The theorem therefore covers cases such as integrated duration processes and processes where the durations divided by their expectation need not be i.i.d. but only strictly stationary, giving added robustness to the empirical results. This result of course only applies to the (1,1) model and it is not easy to extend it in the integrated GARCH case and consequently will not be easy to extend here. However it is a good conjecture that similar results will be true. Engle(1996) presents such a result for the covariance stationary case.

It is important to note that this result does not ensure efficiency of the estimates. Maximum likelihood with the correct density will be the more efficient estimator and it is possible that the gains will be substantial, particularly if samples are not very large. Thus we consider parametric families.

5. Extensions of the Model

The specifications in (11) and (13) can be generalized in many ways. The baseline hazard can be given many parametric shapes. The most popular is to assume that the conditional distribution is Weibull which is equivalent to assuming that x^γ is exponential. Other popular alternatives are the generalized gamma, log logistic, lognormal, as discussed in Lancaster(1990). A further step is to estimate the hazard semiparametrically using methods of piecewise constant, spline, kernel or other smoothers. In the ARCH context Engle and Gonzalez-Rivera(1991) estimated the density using a spline, while in the duration context Engle(1996) used a k-nearest neighbor estimator which is also used here.

In this paper only the Weibull will be estimated. For the Weibull distribution with parameters (κ, γ), the hazard is:

$$(18) \qquad h(x) = \kappa^\gamma x^{\gamma - 1} \gamma$$

and the conditional intensity in terms of ψ_i takes the slightly more complicated form

$$(19) \qquad \lambda\left(t \mid x_{N(t)}, \ldots, x_1\right) = \left(\Gamma\left(1 + \frac{1}{\gamma}\right) \psi_{N(t)+1}^{-1}\right)^\gamma \left(t - t_{N(t)}\right)^{\gamma - 1} \gamma$$

where $\Gamma(\cdot)$ is the gamma function and γ is the Weibull parameter. The conditional intensity is now a two parameter family which can exhibit either increasing or decreasing hazard functions. This makes especially long durations more or less likely than for the exponential depending on whether γ is greater or less that unity respectively. A simple version of the ACD might be specified by (15) and (19) and be called the Weibull ACD(1,1) or WACD(1,1).

The log likelihood for the Weibull ACD is

$$(20) \quad \sum_{i=1}^{N(T)} \ln\left(\frac{\gamma}{x_i}\right) + \gamma \ln\left(\frac{\Gamma\left(1 + \frac{1}{\gamma}\right)x_i}{\psi_i}\right) - \left(\frac{\Gamma\left(1 + \frac{1}{\gamma}\right)x_i}{\psi_i}\right)^{\gamma}$$

for some parameterization of ψ. When $\gamma=1$ this reduces to the log likelihood for the exponential ACD. Clever optimization can avoid repeated evaluation of the gamma function.

The dynamic specification of the conditional duration can also be easily generalized. This could be a non-linear function and it could include other variables such as the marks of previous arrivals as in:

$$(21) \quad \psi_i = \psi(x_{i-1}, x_{i-2}, \ldots, x_{i-p}, \psi_{i-1}, \psi_{i-2}, \ldots, \psi_{i-q}, z_i, z_{i-1}, \ldots, z_{i-1}, \theta)$$

Such a specification includes models analogous to the EGARCH, NGARCH, Power GARCH and many others as possibilities. More interestingly, it allows economic variables to enter the equation which determines the frequency of transactions. From this version of the model one can test hypothesis on economic determinants of the rates of transactions.

A particularly important example is when calendar time effects are important. If the expected duration varies over the time of day then the time at which the duration commences can be thought of as an independent variable. This will be the case for the transactions data to be analyzed in the next section since it is well known that the frequency of transactions is higher near the open and the close of the market. The expected durations therefore can be decomposed into deterministic and stochastic components. Suppose that the deterministic effect of time can be formulated as a multiplicative function so that the detrended or "daily seasonally adjusted" data can be given by

$$(22) \quad \tilde{x}_i = x_i / \phi(t_{i-1}; \theta_\phi)$$

and the expected duration can be written as

$$(23) \quad E_{i-1}(x_i) = \phi(t_{i-1}; \theta_\phi)\psi_i(\tilde{x}_{i-1}, \ldots, \tilde{x}_1; \theta_\psi)$$

then the two sets of parameters can be estimated jointly using the likelihood specified above. In this case ψ_i is interpreted as the expected fraction above or below normal for that time of day.

An alternative approach which was used in Engle and Russell(1995a,1995b) is to initially "daily seasonally adjust" the durations by a regression of x on the spline functions and then model the ratio of actual to fitted value as the daily seasonally adjusted series of durations. This procedure gives very similar results if there are sufficient data since the regression is still a conditional expectation, even though the disturbances are autocorrelated and heteroskedastic:

(24) $E(x_i | t_{i-1}) = \phi(t_{i-1}) + \xi_i$.

6. Modeling the Time Between Transactions of IBM Stock Using the ACD Model.

6.1 Data Description

This section applies the ACD model to IBM transactions data. Initially the focus will be on the time intervals between trades from which the conditional intensity is a high frequency measure of the quantity of transactions. Subsequently models for the price process will be considered.

The data were abstracted from the Trades, Orders Reports and Quotes (TORQ) data set constructed by Joel Hasbrouck and the NYSE. The data set contains detailed information about each transaction occurring on the consolidated market during regular trading hours over a 3 month period beginning November 1, 1990 and ending January 31, 1991[5]. In addition to information about bid and ask quote movements, the volume associated with the transactions, and the transaction prices, there is a time stamp, measured in seconds after midnight, reflecting the time at which the transaction occurred.

---

[5] Some trades occurring during non-trading hours are not recorded. See Hasbrouck Sofianos and Sosebee(1993).

Two days were deleted from the 63 trading days in the 3 month sample. A halt in IBM trading of just over an hour and 15 minutes occurred on Friday, November 23. On December 27th there was a one and a half hour delay in the opening. After deleting these two days there are a total of 58,942 limit and market orders executed at 52,405 unique times. In many cases there are multiple buyers or sellers involved in a trade, and trades can have widely differing numbers of shares transferred. These attributes of a transaction time can be considered marks and can be modeled as such. However in this paper, these marks are not analyzed. See Engle(1996) for an approach to modeling the marks and times jointly and Russell(1996) for a joint model of transaction prices and event arrival times.

The average volume corresponding to each time stamp is 1861 shares. The minimum time between events is 1 second and the maximum duration is 561 seconds or 9 minutes and 21 seconds. The average duration between successive events for IBM (ignoring the overnight waiting time) is 28.38 seconds with a standard deviation of 38.41. Figure 1 is a plot of the histogram for the waiting times. The solid line is the histogram and the dashed line is the exponential distribution with the maximum likelihood estimate of the parameter λ. The waiting times do not appear to be exponential. This does not necessarily mean, however, that the waiting times are not conditionally distributed as an exponential.

The ACD model is proposed as a model for intertemporally correlated event arrival times. To examine the dependence, we calculate the autocorrelations and partial autocorrelations in the waiting times between events. Table 1 contains these autocorrelations and partial autocorrelations after the overnight waiting times are removed. The autocorrelations and partial autocorrelations are far from zero and all the signs are positive. The Ljung-Box statistic is examined to formally test the null hypothesis that the first 15 autocorrelations are 0. The test statistic is distributed as a χ^2_{15} with a 5% critical value of 25. The null is very easily rejected with a chi-squared statistic of 9466 and a corresponding p-value .0000. These long sets of positive autocorrelations are precisely what one finds for autocorrelations of squared returns. Interestingly, volatility clustering and duration clustering exhibit definite similarities which will be developed below.

6.2 Modeling Daily Seasonality and the Open

IBM is traded on several of the US exchanges. These markets are open from 9:30 a.m. EST to 4:00 p.m. EST ; studies have found a seasonal pattern over the course of the trading day[6]. In addition, this paper will present evidence of periods that tend to be relatively active with many traders making transactions and other periods with relatively few transactions. Figure 2 contains a plot of the expected time interval conditioned on the time of day that the duration begins. The expectation is calculated using a nonparametric kernel estimator[7]. The opening is very active with transactions occurring, on average, every 10 seconds. The middle of the day, around 1:00 in the afternoon, tends to have the longest duration between transactions of just over 1 transaction every 35 seconds. Activity picks up again prior to the close with transactions occurring once every 20 seconds, on average.

The extreme short durations around the open are partly a result of the opening auction. During the opening auction, the specialist sets a price to maximize the volume transacted. After the price is set and orders are executed, the transactions must be recorded. Once the open transactions are all recorded, non-opening transactions begin. Because the open is sometimes delayed, the first recorded transaction time of the day varied from the first 34 seconds after 9:30 to later than 9:45. All we observe is the first recorded opening trade and an unknown number of dependent trades that follow. A model of the trading process would be contaminated by including opening trades, hence transactions occurring in the first 20 minutes of the day (9:30-9:50) were deleted. The process for each day is then re-initialized using the average duration over the 10 minutes prior to 10:00. That is, each day starts fresh with the conditional expectation of the first waiting time after 10:00 set equal to the average duration over the 10 minutes prior to

---

[6] See, for example, Jain and Joh (1986), McInish and Wood (1991), or Bollerslev and Domowitz (1993)

[7] The SPLUS subroutine *Supsmu* with cross validation bandwidth selection.

10:00. This design also solves the problem of the carry over of transaction rates from the close one day to the open on the following trading day.

Assuming that the daily seasonal factor $\phi(t_{i-1})$ can be approximated by a cubic spline, nodes were set on each hour. Activity drops off quickly at the end of the day so an additional node for added flexibility was used in the last half hour. The constant in the spline is identified by setting the mean of the predicted daily seasonal factor equal to the observed sample mean.

6.3 Model Estimation

After all the adjustments to the data, there are 46,091 observations on IBM transactions to be analyzed. Maximum likelihood estimation of the ACD model is performed using the Berndt Hall Hall Hausman(1973) algorithm with numerical derivatives. The values used to initialize each day are presented in the appendix. The parameter estimates for the EACD(1,1) and the EACD(2,2) model are presented in table 2. The parameters for the stochastic factor are all significant. The sum of the α_i and β_i is .99623 for the EACD(1,1) model and .99607 for the EACD(2,2) model indicating the process has strong persistence as measured in transaction time.

While virtually all of the parameters for the daily seasonal factor are individually insignificant, a likelihood ratio test rejects the null hypothesis of no seasonal factor. The test statistic has a χ^2_{15} distribution with 5% critical value of 25. The test statistics are 82.38 for the EACD(1,1) and 95.26 for the EACD(2,2). Figure 3 is a plot of the seasonal component for the EACD(2,2) model denoted by the relatively smooth heavy dashed line. Although the scale makes it hard to see, the daily seasonal is approximately an inverted "U" shape with the morning and close exhibiting the highest trading rates and a lull in trading activity just after noon. As expected, this daily seasonal factor is similar in general features to the nonparametric estimate shown in figure 2.

Several characteristics of the data can now be examined. First, consider the intertemporal correlations of the "de-seasonalized" durations given by (22). This series should have a mean of approximately unity and should be free of any daily seasonality. The autocorrelations and partial autocorrelations for this series are presented in table 1.

17

The Ljung-Box statistic associated with these autocorrelations is over 8000 for both the (1,1) and the (2,2) models[8]. This suggests that the large Ljung-Box statistic observed for the raw durations in table 1 is not a result of the seasonal factor alone. This observation is supported by the large t-statistics observed for all models on the parameters α_i and β_i which are designed to capture this intertemporal autocorrelation.

The ACD process assumes that a particular stochastic transformation of the data is i.i.d. Testing this assumption provides a diagnostic check on the model. The "standardized" durations

$$(25) \quad \hat{\varepsilon}_i = \frac{x_i}{\psi_i \phi(t_{i-1})} \ ,$$

are tested for autocorrelation with the Ljung Box statistic with 15 lags. For the EACD(1,1) model, the Ljung-Box is only 38.08 which is much less than the statistic for the raw and seasonally adjusted durations, however it still exceeds the 5% critical value. The Ljung-box for the EACD(2,2), however, reduces the Ljung-Box to 24.9 which is just at the 5% critical value. Of course the independence assumption in equation (8) implies that higher order moments should also be independent. The Ljung-Box statistic for the square of the standardized series associated with the EACD(2,2) is 21.2. Higher order moments are also less than the critical value. These statistics suggest that the model does a good job of accounting for the intertemporal dependence in transaction arrival rates.

A new statistic is presented here to examine non-linear dependence between the standardized durations and the past information set. The seasonally adjusted durations are divided into 23 bins which range from (0,0.1) to (1.9,2.0), and (2,3),(3,4), (4,5) and (5,∞) omitting one bin at 1. The standardized durations are then regressed on indicators for the magnitude of the previous duration; if the standardized durations are indeed i.i.d. then there should be no predictability in this regression. The R^2 of this regression is .003, however because of the 46,000 observations this corresponds to an F statistic of 6.0 which is significant at the .0001 level. The rejection is due to the smallest and largest durations.

---

[8] The test statistic does not vary much across the different distributions to be discussed or the different mean specifications. The test statistic is near 8000 for all the models estimated.

The very smallest duration and the three largest duration cells all have negative t-statistics exceeding 2. This means that the expected durations from the linear model are on average too large after the very shortest durations as well as the longest durations. In a better model, very short durations would have a bigger impact on decreasing expectations, while long durations would have a reduced impact on lengthening durations.

Next, consider the exponential specification of ε. One goodness of fit test is a simple moment condition implied by the exponential distribution. In particular, the exponential distribution implies that the mean should equal the standard deviation. The mean of the standardized durations is unity by first order conditions while the standard deviation is 1.18 for both the EACD(1,1) and EACD(2,2).

A simple test for the null of no excess dispersion is then based on the statistic $\sqrt{N}\left(\dfrac{\hat{\sigma}_{\varepsilon}^2-1}{\sigma_v}\right)$ where $\hat{\sigma}_{\varepsilon}^2$ is the sample variance of $\hat{\varepsilon}$, which should be 1 under the null hypothesis. σ_v is the standard deviation of $(\varepsilon_i-1)^2$ which is equal to $\sqrt{8}$ under the exponential null hypothesis. A straightforward application of a central limit theorem implies this statistic should have a limiting normal distribution under the null with 5% critical value of 1.645. The null of the standard deviation equal to unity is easily rejected with a t-statistics of 30.55 and 30.28 for the EACD(1,1) and EACD(2,2) models respectively. This is strong evidence against the exponential distribution assumption.

Viewing the EACD as a QMLE, we now seek the source of mispecification in ε. Equations (8) and (9) reveal that the empirical distribution of the residual $\left(\dfrac{x_i}{\phi_i\psi_i}\right)$ can be used to obtain estimates of the baseline hazard $\lambda_0(t)$. There are numerous suggestions in the duration literature to calculate the baseline hazard; here a simple semiparametric estimate of the base line hazard using 5% quantiles will be used. Letting e_i denote the empirical residual defining the upper end of the ith 5% quantile the estimator can be written as:

$$(26) \quad h(e_i)=\left(\frac{.05}{1-.05i}\right)\left(\frac{1}{e_i-e_{i-1}}\right)$$

This can be viewed as a kth nearest neighbor semiparametric estimator of the baseline hazard.

Figure 4 presents a plot the semiparametric estimate of the baseline hazard using the EACD(2,2) standardized series. The dashed line represents the exponential hazard and the solid line is the estimate for the empirical hazard. Each dot on the solid line is the estimated hazard for the ith quantile plotted against the midpoint of the ith quantile. These points were then smoothed using a cubic spline. The generally downward sloping hazard is inconsistent with the exponential density. Together, the non-parametric hazard with the parametric conditional mean provide a complete description of the point process.

Alternatively, a parametric hazard can be estimated. The hazard associated with the Weibull distribution is similar in shape to the semi-parametric estimate of the baseline hazard. If γ is less than unity, then the hazard will be monotonically decreasing as seen in Figure 4. Table 3 presents the parameter estimates for the Weibull ACD, or WACD(1,1) and the WACD(2,2) models. The parameter estimate for γ is .9125 and .9126 for the WACD(1,1) and WACD(2,2) models respectively. The Exponential model is easily rejected in favor of the Weibull with a t-statistic associated with the null hypothesis $\gamma = 1$ of approximately 25.5 for both models. The estimates of α_i and β_i are very similar to those of the exponential model. Again, the sum of α_i and β_i is near unity at .9948 and .9953. Although not presented, the graph of the estimated seasonal factor is very similar to that estimated under the Exponential.

Given the similarity in the parameters for the stochastic factor, it is not surprising to find that the standardized series exhibit similar autocovariances and Ljung-Box statistics. If the Weibull specification is correct, then raising the residuals to the power γ should produce a series that is still i.i.d, and is also distributed as a unit exponential. Applying the same set of diagnostics as before to the transformed Weibull standardized series

$$(27) \quad \hat{\varepsilon}_i^\gamma = \left(\frac{x_i}{\psi_i \phi(t_{i-1})} \right)^\gamma .$$

yields Ljung-Box statistics of 39.77 and 25.32 for the WACD(1,1) and WACD(2,2) models respectively. Again these are a great improvement over the Ljung-Box associated with the original series and very similar to those under the exponential distribution. The Ljung-Box for the square of the standardized series is 25.95 for the WACD(2,2) model. The standard deviation of these series has been reduced to 1.07 and 1.06 for the WACD(1,1) and WACD (2,2) models respectively. The test for excess dispersion the of the transformed series in equation (27) yields t-statistics of 11.24 and 11.14 for the WACD(1,1) and WACD(2,2) respectively suggesting there are still problems with the distributional assumptions.

There are several other suggestions in the literature to assess the goodness of fit of the Weibull series[9]. The WACD model implies that the survivor function is given by

$$(28) \quad \overline{F}(x|\psi_i) = \exp\left(-\left(\frac{x}{\phi_i \psi_i}\right)^\gamma\right)$$

It follows that

$$(29) \quad -\log(\overline{F}(x|\psi_i)) = \left(\frac{x}{\phi_i \psi_i}\right)^\gamma$$

so that the negative log of the empirical survivor function should be linearly related to the transformed standardized residuals with a slope of unity. Figure 5 presents this plot. The empirical survivor function has too much weight in the right tail. In particular, there are too many residual values greater than 2.5 or, equivalently, approximately the last decile. This is consistent with the variance tests above that suggested excess dispersion in the residual. Perhaps another parametric distribution would be better suited or we could settle on the semiparametric estimate.

Figure 3 contains a plot of the one-step ahead forecast of the waiting times to the next transaction for a randomly selected day (November 20, 1990). The heavy, relatively smooth dashed line, is the estimated seasonal factor. The heavy solid line is the one-step forecast obtained by taking the product of the stochastic and daily seasonals. The

---

[9] See Kalbfleish and Prentice for a discussion.

observed durations are denoted by the light dashed line. The forecasts exhibit large deviations from the daily seasonal. For example, just after 1:00 the one-step forecast predicts waiting times in excess of 60 seconds while the daily seasonal factor alone is less than 40. Just after 10:00, the forecast drops to near 5 seconds while the seasonal factor alone is more than 30.

7. Some Economics of Market Microstructure

It is natural to ask is why there is stochastic clustering of transaction arrival times since in simple rational expectation and representative agent models, even news will not generate trades. Several explanations are found in the rapidly growing body of literature that examines *market microstructure* of financial markets. See O'Hara(1995) for an excellent survey. This literature often partitions traders into two or more types: informed traders who are assumed to posses information not publicly available, and liquidity traders or those not trading based on a superior information set. Often it is assumed that the composition of the traders is known although individual agents are not identifiable. In a rational expectations setting, the specialist will update the bid and ask quotes in response to order flow characteristics as he infers the new information.

Easley and O'Hara(1992) make the common assumption that liquidity traders arrive randomly according to a Poisson distribution. Informed traders, however, will enter the market only after observing a private, potentially noisy signal. In a rational expectations setting the specialist knows this and will slowly learn of the private information by watching order flow and hence adjust prices accordingly. Informed traders will seek to trade as long as their information has value. Hence, we should see clustering of trading following an information event because of the increased numbers of informed traders.

Admati and Pfleiderer(1988)(1989) develop a model where in addition to informed traders there are two types of uniformed traders. The model supposes a sequence of batch auctions over the course of the trading day with private information lasting a single period.

The "discretionary" liquidity traders have some choice over the time at which they transact, while the "non-discretionary" liquidity traders are again assumed to arrive in a random fashion. Because the bid ask spread in the model is inversely related to the discretionary trader volume, it is optimal for discretionary liquidity traders to lump their trades together in the same batch auction. The informed traders would also choose to trade at this time but the number of informed traders is exogenous so the increases in volume are associated with increased numbers of liquidity traders.

To investigate whether the episodes of high transaction intensity are due to liquidity traders or informed traders, it is necessary to examine prices and bid ask spreads as well as transaction times which we do in the next section. In the rational expectations model of Admati and Pfleiderer, there is a constant amount of private information revealed each period so the variability of the price should be independent of volume. In Easley and O'Hara, increased numbers of transactions are due to information events and the naturally increased numbers of informed traders. As a consequence, bid ask spreads widen and price volatility increases.

8. An ACD Model for Price Movements and Some Tests of Market Microstructure

The models estimated above defined an event arrival using transaction arrival times. Alternatively, events can be defined by a subset of the transaction arrival times with specific characteristics or marks. For example, to focus on the price process, consider selecting only the points for which the price has changed. An event arrival time is now characterized not just by a transaction occurring, but a transaction occurring at a new price.

The point process literature calls the price event arrival times a *thinned point process*. Because the price marks themselves are presumably autocorrelated the process of selecting the subsample is called *dependent thinning*. While theorems exist relating the intensity function of the thinned point process to that of the original sample they generally assume that the original sample is a time homogeneous Poisson process.

The ACD model can now be applied to these new event arrival times. Clearly, this no longer models the arrival rate of traders but rather how quickly the price is changing. The intensity function is now a measure of the instantaneous probability of a price movement called a "price intensity". Similarly, the waiting times between these events are called "price durations". This is intuitively a model for the inverse of volatility; in the next section this relation is made precise.

Define the midpoint of the bid-ask spread or "midprice" to be the current price. This eliminates the problem of "bid-ask bounce" although price discreteness remains a complication. Define the midprice at time t_i as

$$(30) \quad p_i = \frac{bid_i + ask_i}{2}.$$

Bid_i and ask_i are the current bid and ask prices associated with transaction time t_i. Examining the IBM quotes reveals that sometimes a bid or ask quote will move by an 1/8th for only a few transactions and then return to its previous level. This might not be considered a price change as it might be a recording or quoting error, or just reflect inventory control or a single particularly large transaction that temporally moves the price. To avoid counting these minor, potentially transitory price movements, define a price movement as any movement in the midprice greater than or equal to some constant c. In addition, an attempt was made to make the thinning robust to errant quotes by requiring two consecutive midprice values to have moved by at least the threshold value.

The median spread for IBM is .25, or 2 ticks. Occasionally the spread is more than a dollar, and the minimum recorded spread is 1 tick at .125. A value of c = .25 is used so that altogether the bid or ask quotes must move by a total of 4 ticks to initiate a new price duration.

The sample size is reduced to 1347 after thinning or about 3% of the original sample. The mean price duration is just under 15 minutes at 860 seconds and there is an average of 25.5 price events per day. The minimum price duration is 1 second while the maximum is almost 3 hours at 10,609 seconds. There is excess dispersion relative to the exponential with a standard deviation 1.43 times the mean. The Ljung-Box statistic

associated with the price durations is 72. All autocorrelations are positive indicating clustering of price events.

Volatility is also known to exhibit intradaily patterns so the daily seasonal factor model specification with splines was used again. For many of the days no price events were recorded before 10:00. The initial value for these days was set equal to 1.0. A list of the initial conditions for the price event model is given in the appendix.

The (2,2) model specification is required to whiten the standardized series. Estimates of the EACD(2,2) model and the WACD(2,2) model are presented in table 4. The Ljung-Box statistic associated with the de-seasonalized series increases slightly to 104. The sum of α_i and β_i is .9191 and .9043 for the exponential and Weibull models respectively. This is much less than for the transaction event models. It is difficult however to compare the persistence in the transaction and price models because the transaction durations are much shorter on average than the price durations. The Weibull parameter is again significantly less than 1 indicating a downward sloping hazard.

Although none of the daily seasonal factors are individually significant, the value of the likelihood ratio test for the EACD(2,2) model is 157.9 with a critical value of 25 suggesting the seasonal factor should be included. Plots of the daily seasonal factors for the EACD(2,2) and WACD(2,2) models are presented in figure 6. As expected, the price durations are shortest in the morning and just prior to the close with a noticeable lull between 12:00 and 1:00.

The Ljung-Box has been reduced to approximately 7.05 for the exponential and 6.35 for Weibull. Again there is evidence of remaining excess dispersion for the exponential model with a test statistic of 7.93. Figure 7 gives the empirical estimate of the hazard which is essentially monotonic and downward sloping which is consistent with the Weibull hazard function. This time the moment condition for excess dispersion is not rejected for the Weibull standardized series with a test statistic of -.14. Figure 8 presents the log-linear plot of the empirical survivor function against the standardized duration. Again, the solid line should be approximately linear with slope equal to unity (represented by the dashed line) if the true distribution were Weibull. The plot also suggests that the Weibull distribution fits the price duration data well.

Before testing the market microstructure hypotheses, we examine the relationship of the price intensities to standard measures of volatility.

8.1 A Relationship Between the Price Intensity and Volatility.

From equation (1), the price intensity function characterizes the probability that a price event will occur in the next instant, conditional on the history of price events. More precisely, the probability that the price changes by c in an interval Δt is $\lambda(t|t_{i-1},...,t_0)\Delta t + o(\Delta t)$, and otherwise there is no change. This can be viewed as a binomial price process where the time between events is random as well as the direction of change. Let the instantaneous volatility be defined as:

$$(31) \quad \sigma^2(t) = \lim_{\Delta t \to 0} E\left\{ \frac{1}{\Delta t}\left[\frac{P(t+\Delta t)-P(t)}{P(t)} \right]^2 \right\}$$

Substituting and taking limits, the smaller terms vanish yielding the expectation of the conditional volatility per second in the next instant:

$$(32) \quad \sigma^2\left(t|t_{N(t)},...,t_1\right) = \left(\frac{c}{P(t)} \right)^2 \lambda\left(t|t_{N(t)},...,t_1\right)$$

Since the intensity function is defined for all $t>t_0$ we have a continuous picture of volatility over the trading day.

A plot of the instantaneous volatility over three consecutive days (January 21, 22, and 23) is given in the top half of figure 9. These days were particularly volatile so they might be interesting to examine. The intensities were constructed using the EACD(2,2) model with nonparametric hazard. Because the units of volatility per second are not very intuitive the volatility was annualized as if the volatility per second were constant for all trading hours over the course of a year[10]. The spikes in the function indicate event arrival times. The downward sloping nonparametric hazard function, that was presented in figure 7 implies that we are most likely to observe a price event immediately following a price

---

[10] (252 trading days/year)*(6.5 hours/day)*(3600 seconds/hour)=5,896,800 trading seconds/year

event. Focusing on the general level rather than the steep descent and upward spike we see different patterns over the course of each day. The first day plotted (910121) exhibits high volatility in the morning and just prior to close. The volatility for day 2 (910122) does not exhibit the large swings and clumping of volatility that day 1 does. Day 3 (910123) has four bursts of volatility over the course of the day: two in the morning, one around mid-day and finally, another burst near the close.

At each point in time a volatility forecast can be computed. Although this can be computed analytically in event time, a closed form solution does not exist in calendar time but the forecast can be computed by simulation. These forecasts will not have the spikes that the instantaneous expectations have since the conditioning information becomes less valuable as the forecast horizon increases.

8.2 Testing Market Microstructure Hypotheses About Transaction Clustering.

The bottom of figure 9 contains a plot of the transaction intensity over the course of the same three days used for the instantaneous volatility. The hazards were also calculated using the EACD(2,2) model with nonparametric hazards. Casual examination suggests a correlation between the two intensity functions but it is far from perfect as there are episodes of high transaction intensity without price movement.

The Easley and O'Hara model predicts that the number of transactions would influence the price process through information based clustering of transactions. The Admati and Pfleiderer model predicts that the number of transactions would have no impact on the price intensity. While casual evidence may be ambiguous, these hypotheses can now be statistically tested.

A measure of the intensity of trading is constructed using the number of transactions per second over each price duration. The daily seasonality was then removed by dividing each value by it's conditional mean calculated using the same spline function as used for the transaction duration models. This procedure introduces no biases into the test statistic as it recognizes that the seasonal pattern of transactions and prices may not be the same. The seasonal splines in the price durations are estimated jointly as before.

The price duration EACD(2,2) model is used as the base model with robust standard errors to insulate the results from the validity of the Weibull parameterization. Model 1 in table 5 includes the past transaction rate which is statistically significantly negative with a robust t-statistic of -12.65. The negative sign means that the expected price durations are shorter, and equivalently the volatility is higher, following periods of high transaction rates. This is consistent with the Easley and O'Hara model.

Further evidence can be found using volume per transaction and spread as these are also thought to carry information about the existence of private information. For a survey of the volume price relationship see Karpoff(1987) or Stickel and Verrecchia(1993). Starting with Glosten and Milgrom(1985), asymmetric information models generally predict that spreads tend to become wider as the probability increases of informed agent trading. For an interesting empirical investigation see Hasbrouck(1991). Model 2 adds the average volume per transaction observed over the previous price duration and the percent bid ask spread at the end of the last duration, each corrected for daily seasonality using the same procedure as before. The coefficient on the number of transactions per second is reduced slightly but is still very significant and negative. The spread and volume are also highly significant with robust t-statistics of -15.68 and -4.5. These strong results further support the Gosten and Milgrom as well as Easley and O'Hara conjecture that information based trading predicts higher volatilities.

Finally, consider the possibility that clustering of trading may be occurring at different times for different reasons. Perhaps there are times that transaction clustering is due to information based trading and other times transaction clustering is due to liquidity based trading. If the specialist can distinguish these periods, he will set narrow spreads when the fraction of informed traders is small and wide spreads when the fraction of informed traders is large. This suggests using the spread to break up liquidity clustering and information based clustering.

Model 3 allows the effect of the number of transactions in a "wide spread" state to be different from the effects in a "narrow spread" state. If the heterogeneous effect is correct we would expect the coefficient on the "wide spread" number of transactions to be negative reflecting more rapid price movements when the fraction of informed traders is

high.   The coefficient on the number of transactions per second in the "narrow spread" state should be insignificant according to Admatti and Pfleiderer. The Admatti Pfleiderer model is formulated in an environment where informed traders always exist. Interpreting the Admatti Pfleiderer results in a more general context where private information does not always exist, we might expect the coefficient on the "narrow spread" state to be positive reflecting a very liquid market. That is, in absence of informed traders, the market is probably more liquid when trading activity is higher.

The coefficient on the number of transactions per second in the "wide spread" state is -.0365 with a robust t-statistic of -17.89 while the coefficient in the "small spread" state is .0026 with a robust t-statistic of .12. If the spread is a good measure of the fraction of informed traders, then these results suggest the effect of transaction rates on volatility varies as a function of the fraction of traders that are informed. The effect of the transaction frequency under narrow spreads is statistically insignificant so Model 4 contains the volume, spread, and transaction frequency in only the "wide spread" state.

These results shed light on the causes of transactions clustering. The price tends to move quickly following high transaction rates if the spread is relatively wide which is interpreted as when informed traders are likely to be active. The price tends to move less quickly or is perhaps unaffected following higher transaction rates when spreads are narrow which is when liquidity traders are inferred to be dominant. This suggests that both liquidity and information based clustering of transaction rates occurs.

9. Conclusions

This paper proposed a new technique for modeling irregularly spaced time series data using a new class of self-exciting point process models that seem particularly well suited for financial data. The model can be interpreted in the class of time deformation models that are becoming popular in finance. Asymptotic results are extracted from the ARCH-GARCH literature to obtain QMLE results for the exponential version of the ACD model.

The ACD model did a very good job of reducing excessively large Ljung-Box statistics associated with waiting times between transaction events to statistically insignificant, or marginally significant levels. Excess dispersion tests suggest that neither the Weibull nor the exponential versions of the ACD model may be fully appropriate parameterizations for transaction event arrival times. These models are presented as a useful starting point along with diagnostics and robust standard errors.

A model for the price process is estimated by selecting a subsample of the data which carries particular information about price movements. The Weibull model appears to fit this distribution well as it easily passes both the intertemporal dependence and excess dispersion diagnostic tests. This model can be interpreted as a model for instantaneous volatility. It is interesting because the model provides a continuous picture of volatility over the course of the day. The price model is used to test hypotheses about the source of clustering of transaction times. Evidence is presented that clustering of transactions occur both due to bunching of informed traders and, when spreads are small, also to clustering of liquidity traders.

The avenues for future research stretch in several directions. Recognizing that the underlying price process specified in Section 8 is essentially a binomial tree with random spacings, implications for derivative pricing and term structure of interest rates seem direct. Joint modeling of volumes, transaction prices and quotes would give better understanding of the fundamental mechanisms of NYSE markets including time variation in liquidity. Additional insight will come from joint modeling of more than one asset where the second could be a related stock or a derivative with duration spillovers.

Table 1. Autocorrelations and Partial Autocorrelations of Trading Intervals

	Raw Durations		De-Seasonalized Durations	
	acf	pacf	acf	pacf
lag 1	.155	.155	.144	.144
lag 2	.132	.110	.122	.103
lag 3	.126	.094	.113	.085
lag 4	.125	.086	.110	.076
lag 5	.122	.075	.111	.072
lag 6	.108	.056	.098	.053
lag 7	.118	.065	.109	.063
lag 8	.123	.066	.113	.063
lag 9	.116	.054	.102	.047
lag 10	.104	.038	.095	.038
lag 11	.106	.041	.098	.042
lag 12	.106	.039	.095	.036
lag 13	.105	.037	.093	.033
lag 14	.099	.030	.089	.029
lag 15	.097	.028	.088	.028

Ljung-Box(15) = 9466.43	Ljung-Box(15) = 7807.41
Sample Size=46091	Sample Size=46091

Table 2: Parameter Estimates for the EACD Transaction Models

	EACD(1,1)			EACD(2,2)		
	Estimate	t-Statistic	Robust t-Statistic	Estimate	t-Statistic	Robust t-Statistic
ω	.00568	11.94	9.16	.00584	9.17	6.94
α_1	.06306	41.61	33.70	.08295	19.86	15.30
α_2				-.01608	-2.75	-2.05
β_1	.93317	574.66	447.31	.79948	12.07	9.44
β_2				.12971	2.09	1.63
c_{11}	28.402	-	-	26.493	-	-
d_{11}	-8.283	-.80	-.59	-1.962	-.19	-.14
d_{21}	17.365	.81	.58	10.307	.49	.36
d_{22}	-21.502	-1.55	-1.14	-20.944	-1.54	-1.12
d_{23}	-8.167	-.56	-.41	-6.102	-.42	-.31
d_{24}	-39.286	-2.62	-1.74	-37.202	-2.52	-1.68
d_{25}	-25.132	-1.77	-1.35	-26.129	-1.88	-1.42
d_{26}	-39.524	-1.13	-.8	-46.359	-1.33	-.95
d_{27}	-29.110	-.61	-.44	-32.463	-.68	-.49
d_{31}	-7.222	-.59	-.42	-4.520	-.38	-.27
d_{32}	16.168	1.81	1.33	15.487	1.76	1.28
d_{33}	5.274	.55	.39	3.631	.38	.27
d_{34}	23.165	2.41	1.64	22.426	2.38	1.61
d_{35}	17.604	1.89	1.45	18.267	2.00	1.52
d_{36}	44.198	.92	.66	54.330	1.13	.81
d_{37}	51.114	.72	.48	54.699	.77	.51

Statistics from Residuals

Mean	.9999		Mean	.9994
Std Dev	1.1843		Std Dev	1.1828
Ljung-Box	38.0864		Ljung-Box	24.9446
Excess Disp.			Excess Disp.	
Test Statistic	30.55		Test Statistic	30.28

$$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j \tilde{x}_{i-1} + \sum_{j=1}^{q} \beta_j \psi_{i-1} \text{ where } \tilde{x}_{i-1} \equiv \frac{x_{i-1}}{\phi(t_{i-1})}$$

$$\phi(t_{i-1}) = \sum_{j=1}^{K} I_j \left[c_j + d_{1,j}(t_{i-1} - k_{j-1}) + d_{2,j}(t_{i-1} - k_{j-1})^2 + d_{3,j}(t_{i-1} - k_{j-1})^3 \right]$$

where I_j is the indicator variable for the jth segment of the spline $\{I_j=1 \text{ if } k_{j-1} \leq t_{i-1} < k_j, =0 \text{ otherwise}\}$.

For $j>1$ $c_{i,j}$ and $d_{1,j}$ are restricted by the usual differentiability conditions.

c_{11} is normalized by restricting the unconditional mean of the seasonal factor to equal the observed sample mean.

Table 3: Parameter Estimates for the WACD Transaction Models

	WACD(1,1)			WACD(2,2)		
	Estimate	t-Statistic		Estimate	t-Statistic	
ω	.00592	10.20		.00616	7.68	
α_1	.06262	34.45		.08321	16.56	
α_2				-.01601	-2.21	
β_1	.93221	469.57		.79768	9.59	
β_2				.13046	1.67	
γ	.91250	19.28	(Ho: γ=1)	.91263	24.82	(Ho: γ=1)
c_{11}	27.142	-		25.682	-	
d_{11}	-5.163	-.42		-.603	-.05	
d_{21}	13.506	.53		8.529	.34	
d_{22}	-19.359	-1.16		-19.866	-1.21	
d_{23}	-4.574	-.26		-3.949	-.22	
d_{24}	-36.359	-1.99		-35.370	-1.97	
d_{25}	-25.047	-1.45		-26.041	-1.54	
d_{26}	-46.955	-1.08		-51.634	-1.20	
d_{27}	-35.845	-.60		-37.643	-.63	
d_{31}	-5.691	-.39		-3.69	-.26	
d_{32}	14.493	1.34		14.545	1.36	
d_{33}	2.572	.22		1.988	.17	
d_{34}	21.818	1.87		21.562	1.88	
d_{35}	17.826	1.57		18.465	1.65	
d_{36}	54.853	.92		61.749	1.04	
d_{37}	59.144	.66		61.186	.69	

Statistics from Residuals

Mean	1.0002		Mean	1.0003	
Std Dev	1.0715		Std Dev	1.0709	
Ljung-Box	39.7715		Ljung-Box	25.3202	
Excess Disp.			Excess Disp.		
Test Statistic	11.24		Test Statistic	11.14	

$$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j \tilde{x}_{i-1} + \sum_{j=1}^{q} \beta_j \psi_{i-1} \text{ where } \tilde{x}_{i-1} \equiv \frac{x_{i-1}}{\phi(t_{i-1})}$$

$$\phi(t_{i-1}) = \sum_{j=1}^{K} I_j \left[c_j + d_{1,j}(t_{i-1} - k_{j-1}) + d_{2,j}(t_{i-1} - k_{j-1})^2 + d_{3,j}(t_{i-1} - k_{j-1})^3 \right]$$

where I_j is the indicator variable for the jth segment of the spline $\{I_j=1$ if $k_{j-1} \leq t_{i-1} < k_j, =0$ otherwise$\}$.

For $j>1$ $c_{i,j}$ and $d_{1,j}$ are restricted by the usual differentiability conditions.

c_{11} is normalized by restricting the unconditional mean of the seasonal factor to equal the observed sample mean.

Table 4: Parameter Estimates for the Price Models

	EACD(2,2)			WACD(2,2)		
	Estimate	t-Statistic	Robust t-Statistic	Estimate	t-Statistic	
ω	.1150	4.95	3.13	.1250	3.31	
α_1	.0703	3.66	2.30	.0745	2.36	
α_2	.1983	7.59	5.96	.2180	4.94	
β_1	.1520	2.51	1.55	.1246	1.33	
β_2	.4985	8.76	5.28	.4872	5.53	
γ				.7650	12.40	$(H_o\ \gamma=1)$
c_{11}	794.739	-	-	763.080	-	-
d_{11}	476.460	.42	.25	459.687	.26	
d_{21}	-195.196	-.08	-.04	-101.911	-.02	
d_{22}	275.513	.21	.10	261.194	.12	
d_{23}	238.016	.15	.08	524.297	.21	
d_{24}	-609.202	-.37	-.24	-668.336	-.24	
d_{25}	16.744	.01	.01	86.048	.03	
d_{26}	-264.385	-.09	-.06	-54.562	-.01	
d_{27}	-3379.156	-.90	-1.15	-3807.942	-.57	
d_{31}	-69.153	-.05	-.02	-127.715	-.06	
d_{32}	-67.975	-.08	-.03	-76.041	-.05	
d_{33}	-319.247	-.29	-.18	-485.941	-.28	
d_{34}	437.515	.43	.28	478.477	.28	
d_{35}	-106.821	-.12	-.09	-180.527	-.12	
d_{36}	1176.874	.29	.22	1089.532	.15	
d_{37}	2788.492	.54	.77	3031.447	.33	

Statistics from Residuals

Mean	.9999		Mean	.9999
Std Dev	1.2696		Std Dev	.9943
Ljung-Box	7.05		Ljung-Box	6.35
Excess Disp.			Excess Disp.	
Test Statistic	7.93		Test Statistic	-.14

$$\psi_i = \omega + \sum_{j=1}^{p}\alpha_j\tilde{x}_{i-1} + \sum_{j=1}^{q}\beta_j\psi_{i-1} \text{ where } \tilde{x}_{i-1} \equiv \frac{x_{i-1}}{\phi(t_{i-1})}$$

$$\phi(t_{i-1}) = \sum_{j=1}^{K} I_j\left[c_j + d_{1,j}(t_{i-1}-k_{j-1}) + d_{2,j}(t_{i-1}-k_{j-1})^2 + d_{3,j}(t_{i-1}-k_{j-1})^3\right]$$

where I_j is the indicator variable for the jth segment of the spline $\{I_j=1 \text{ if } k_{j-1}\leq t_{i-1}<k_j, =0 \text{ otherwise}\}$.

For $j>1$ $c_{i,j}$ and $d_{1,j}$ are restricted by the usual differentiability conditions.

c_{11} is normalized by restricting the unconditional mean of the seasonal factor to equal the observed sample mean.

Table 5: Price Model Estimates with Robust t-Statistics

Parameter	Model 1	Model 2	Model 3	Model 4
ω	.2107 (6.14)	.3027 (18.22)	.2630 (12.79)	.2647 (22.69)
α_1	.0457 (2.60)	.0507 (2.24)	.0527 (2.38)	.0531 (2.74)
α_2	.1731 (5.94)	.1578 (5.19)	.1492 (4.77)	.1503 (4.82)
β_1	.0769 (1.00)	.1646 (1.61)	.2007 (1.99)	.1968 (2.22)
β_2	.5609 (8.07)	.4600 (5.16)	.4409 (4.76)	.4427 (5.11)
#Trans/Sec	-.0440 (-12.65)	-.0359 (-13.40)		
Spread		-.0782 (-15.68)	-.0741 (-21.58)	-.0740 (-21.19)
Volume/Trans		-.0041 (-4.58)	-.0043 (-9.14)	-.0043 (-9.67)
#Trans/Sec (Wide Spread)			-.0365 (-17.89)	-.0364 (-18.44)
#Trans/Sec (Narrow Spread)			.0026 (.12)	

Figure 1: Histogram of Transaction Durations

Figure 2: Nonparametric Estimate of Seasonal Pattern for Transaction Durations

**Figure 3: Estimate of Seasonal Factor and One-Step forecast for
Transaction Durations**

Figure 4: Semiparametric Estimate of Baseline Hazard for Transaction Durations

Figure 5: Plot of -Log Empirical Survivor for Weibull Transaction Model

Figure 6: Estimate of Seasonal Factor for Price Durations

Figure 7: Semiparametric Estimate of Baseline Hazard for Price Durations

Figure 8: -Log Empirical Survivor for Weibull Price Model

Figure 9: Forecast of Instantaneous Volatility(Top) and Transaction Intensity (Bottom)

Appendix
Values Used to Initialize ACD Process for Each Day

Date	Transaction Model	Price Model	Date	Transaction Model	Price Model
901101	1.276	1.0	901218	.810	1.190
901102	1.093	1.0	901219	.927	1.0
901105	1.410	1.0	901220	.536	.434
901106	.648	.530	901221	.992	1.0
901107	1.456	1.0	901224	1.208	.588
901108	1.915	1.44	901226	6.471	.880
901109	1.056	1.0	901228	1.284	.416
901112	.563	1.0	901231	1.691	1.0
901113	.529	.273	910102	.575	.483
901114	.595	1.0	910103	.656	1.162
901115	.871	1.0	910104	1.121	1.0
901116	.861	4.715	910107	.940	1.0
901119	1.005	1.0	910108	.589	.633
901120	1.315	1.0	910109	.413	.413
901121	.846	1.0	910110	.563	1.0
901126	.900	3.269	910111	1.203	1.0
901127	1.557	1.0	910114	.948	.277
901128	1.058	.537	910115	.921	1.0
901129	.944	.421	910116	1.300	.574
901130	.859	1.0	910117	.199	.155
901203	1.545	2.138	910118	.487	1.0
901204	1.408	1.0	910121	.586	.037
901205	.763	1.687	910122	.498	.392
901206	.335	2.281	910123	.7531	1.0
901207	.511	.268	910124	.868	1.0
901210	.712	1.0	910125	.502	.346
901211	.976	1.0	910128	.240	.152
901212	.968	1.0	910129	.578	1.0
901213	.772	2.242	910130	1.197	1.456
901214	1.442	1.0	910131	.593	1.0
901217	1.136	.588			

References

1. Admati, Anat R. and Paul Pfleiderer, 1988, A theory of Intraday Patterns: Volume and Price Variability, *The Review of Financial Studies* 1, 3-40.

2. Berndt, E., B. Hall, R. Hall, J. Hausman,(1974), Estimation and Inference in Nonlinear Structural Models, *Annals of Economic and Social Measurement*, 3,pp 653-665

3. Bera, Anil and Mathew Higgins, 1992, Survey of ARCH Models: Properties, Estimation and Testing, unpublished manuscript, Department of Economics, University of Illinois

4. Bollerslev, Tim, 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327

5. Bollerslev, Tim, R. Chou, and K. Kroner, 1992 ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics*, 52, 5-59

6. Bollerslev, Tim and Ian Domowitz, 1993, Trading Patterns and Prices in the Interbank Foreign Exchange Market, *The Journal of Finance*, 48, 1421-1443

7. Bollerslev, Tim, Robert Engle, and Dan Nelson, 1994, ARCH models, *in Handbook of Econometrics IV*, (Robert Engle and Dan McFadden eds.), North Holland

8. Clark, Peter, 1973, A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, *Econometrica*, 41, 135-156

9. Cox, D.R., 1955, Some Statistical Models Connected with Series of Events (with Discussion), *Journal of the Royal Statistical Society Ser. B*, 17, 129-164

10. Cox, D. R., 1972a, Regression models and life tables (with discussion*). Journal of the Royal Statistical Society Ser. B*, 34, 187-220

11. Cox, D. R., 1972b, The Statistical Analysis of Dependencies in Point Processes, in *Symposium on Point Processes*, (P.A.W. Lewis, ed.), New York, John Wiley 55-66

12. Easley and O'Hara, 1992, Time and the Process of Security Price Adjustment. *The Journal of Finance* 19, 69-90

13. Engle, Robert, 1982, Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50, 987-1008

14. Engle, Robert, 1996, The Econometrics of Ultra-High Frequency Data, UCSD Department of Economics Working Paper 96-15

15. Engle, Robert and Gloria Gonzalez-Rivera, 1991, Semiparametric ARCH Models, *Journal of Business and Economic Statistics*, 9, 345-359

16. Engle, Robert and J. Russell, 1995a, Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model, Forthcoming in *Journal of Empirical Finance*

17. Engle, Robert and J. Russell, 1995b, Autoregressive Conditional Duration: A New Model for Irregularly Spaced Data, University of California, San Diego, unpublished manuscript

18. Gaver, D.P. and P.A.W Lewis., 1980, First Order Autoregressive Gamma Sequences and Point Processes, *Advances in Applied Probability* 12, 727-745

19. Ghysels, Eric and J. Jasiak 1994, Stochastic Volatility and Time Deformation: An Application to Trading Volume and Leverage Effects, C.R.D.E., Universite de Montreal, unpublished manuscript

20. Glosten, Lawrence R., Milgrom P., 1985, Bid Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents, *Journal of Financial Economics* 14 71-100

21. Grandell, J, 1976, *Doubly Stochastic Poisson Processes* (Lecture Notes Math529), Springer Verlag, New York

22. Hasbrouck, J., 1991, Measuring the Information Content of Stock Trades, *The Journal of Finance* 66,1, 179-207

23. Hasbrouck, J., Sofianos, and Sosebee, 1993, New York Stock Exchange Systems and Trading Procedures, NYSE Working Paper #93-01

24. Hawkes, A.G., 1971a, Spectra of Some Self-Exciting and Mutually Exciting Point Processes, *Biometrika* 58, 83-90

25. Hawkes, A.G., 1971b, Point Spectra of Some Mutually Exciting Point Processes, *Journal of the Royal Statistical Society*, 33, 438-443

26. Hawkes, A.G., 1972, Spectra of Some Mutually Exciting Point Processes with Associated Variables, *Stochastic Point Processes* (PAW Lewis, ed.), Wiley, New York, 261-271

27. Heckman, J. J., and Borjas, B. J., 1980, Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence. *Economica* 47: 247-283

28. Heckman, J., 1981, Heterogeneity and State Dependence, in S. Rosen, ed. *Studies in Labor Markets*, National Bureau of Economic Research pp91-139

29. Jain, P. J., and Joh, G., 1988, The Dependence Between Hourly Prices and Trading Volume, *Journal of Financial and Quantitative Analysis* 23, 269-284.

30. Jacobs,P.A. and P.A.W.Lewis, 1977 A Mixed Autoregressive-Moving Average Exponential Sequence and Point Process(EARMA1,1), *Advances in Applied Probability*, 9,pp87-94

31. Jones, C., and G. Kaul, 1994, Transactions Volume and Volatility, *Review of Financial Studies* 7, 631-651.

32. Kyle, Albert, 1985, Continuous Time Auctions and Insider Trading, *Econometrica* 53, 1315-1336

33. Kalbfleisch, J., and R. Prentice, 1980, *The Statistical Analysis of Failure Time Data*, John Wiley & Sons.

34. Karpoff, J., 1987, The Relation Between Price Change and Trading Volume: A Survey, *Journal of Financial and Quantitative Analysis* 22, 109-126

35. Lancaster, Tony, 1979, Econometric Methods for the Duration of Unemployment, *Econometrica*, 47,4:pp 939-956

36. Lancaster, Tony, 1990, *The Econometric Analysis of Transition Data,* Econometric Society Monographs, Cambridge University Press

37. Last, G., and A. Brandt, 1995, *Marked Point Processes on the Real Line The Dynamic Approach*, Springer

38. Lawrance,A.J. and P.A.W. Lewis 1980, The Exponential Autoregressive-Moving Average EARMA(p,q) Process, *Journal of the Royal Statistical Society*, B, 42, pp. 150-161

39. Lee, S., and B. Hansen, 1994, Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator, *Econometric Theory* 10, 29-52

40. Lumsdaine, R., 1996, Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models, *Econometrica* 64, 575-596.

41. McInish T., and R. Wood, 1991, An Analysis of Intradaily Patterns in Bid/Ask Spreads for NYSE Stocks, Working Paper, Memphis State University.

42. O'Hara, M., 1995, *Market Microstructure Theory*, Basil Blackwell Inc.

43. Muller U.A., M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, C. Morgenegg, Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law, and Intraday Analysis, *Journal of Banking and Finance*, 14, 1189-1208

44. Rubin, I. 1972a, Regular Point Processes and their Detection, *IEEE Trans. on Information Theory*, ITT-18, 547-557

45. Russell, Jeffrey, Econometric Analysis of Irregularly Spaced Transactions Data Using a New Class of Accelerated Failure Time Models with Applications to Financial Transaction Data. Department of Economics University of California, San Diego Ph.D Dissertation.

46. Snyder, D.L. and M.I. Miller, 1991,*Random Point Processes in Time and Space*, Second Edition, Springer-Verlag

47. Stickel, S, and R. Verrecchia, 1993, Evidence that Volume Sustains Price Changes, Working Paper, The Wharton School.

48. Ogata, Y., and K. Katsura, 1986, Point-Process Models with Linearly Parameterized Intensity for the Application of Earthquake Catalogue, *Journal of Applied Probability* 23A, 231-240.

49. Stock James, 1988, Estimating Continuous time Processes Subject to Time Deformation, *Journal of the American Statistical Association*, 83, 77-85

50. Tauchen, George and Mark Pitts, 1983, The Price Variability-Volume Relationship on Speculative Markets, *Econometrica*, 51, 485-505

51. Vere-Jones, D. and T. Ozaki, 1982, Some Examples of Statistical Inference Applied to Earthquake Data, *Annals of the Institute for Statistical Mathematics* 34, 189-207

52. Wold, H. 1948, On Stationary Point Processes and Markov Chains. *Skand. Aktuar.*, 31, 229-240