Modelling the Dynamics of Financial Transactions Data

William T.M. Dunsmuir

UNSW

August 8 2009
Objective: Review models for the dynamics of transaction by transaction stock price data

Outline:

- IBM Example
- Market microstructure hypotheses.
- Ultra-High Frequency Data.
- What are these models used for?
- Types of observation driven models - ADS, ACD, Multinominal.
- Pre-processing the data - adjustment for time of day effects.
- Assessing the constancy of model parameters from day to day.
- The general GAS model.
Example IBM Jan 1 to Feb 28, 2002

IBM Average Price ($) : Jan 2 - Feb 28 2002

IBM Price Changes (1c Ticks) : Jan 2 - Feb 28 2002

IBM Time Between Trades (seconds): Jan 2 - Feb 28 2002
Russell and Engle (2005): "These new data sets provide us with an unprecedented microscopic view of the structure of financial markets that was previously impossible with time aggregated data."

Data are fundamentally irregularly spaced in time. Need models and methods:

- directly tailored to irregularly spaced data.
- for integer valued outcomes.
Random Variables of Interest

Observed:

- Time of transaction, or durations between successive transactions
- Prices, or changes in prices (within days, not between days).
- Marks: vector observed at time of transaction (volume and price of contract, posted bid and ask prices ...)

Derived:

- Active Trade - Price changed or not.
- Direction of Trade - Up or DownPrice.
- Size of price change.
Want an econometric model for the joint process of transaction arrivals and tick-by-tick price movements.

Let:

- $t_i = \text{Arrival time of } i\text{th transaction}$
- $y_i = \text{trade-to-trade change in asset price at time } t_i$
- $x_i = t_i - t_{i-1} = \text{duration between trades}$

Objective: model the joint distribution of durations and price changes given the past (and other covariates $z^{(i-1)}$)

\[
f(y_i, x_i | y^{(i-1)}, x^{(i-1)}, z^{(i-1)}) = g(y_i | y^{(i-1)}, x^{(i)}, z^{(i-1)}) \times q(x_i | y^{(i-1)}, x^{(i-1)}, z^{(i)})
\]
The main reasons for building these models are to:

1. Characterize and explain the intraday process for transaction arrivals and price changes.
2. Forecast these.
3. Provide empirical assessment and testing of various market microstructure hypotheses about the relationship between price movements and other marks of the trading process.
How Can These Models be Used to Test Market Microstructure?

In order to use the econometric model to test market microstructure hypotheses other information in addition to lagged durations such as:

- volume
- bid-ask spread
- volatility of the returns

will need to be included in the model.

Market microstructure theory is assessed using the signs, directions and lags of these model terms.
Two types of models proposed for the duration between successive trades:

1. ACD models (Engel)
2. Discrete time BIN model (Rydberg and Shephard)
The essential idea is to model the durations conditional on the past with:

1. a unit mean distribution (e.g. generalized gamma) for the excess durations (durations divided by their expected value)
2. an autoregressive moving average (ARMA) with explanatory variables for the evolution of the expected durations.
ACD model - Evolution of Mean Durations

Let \( \mathcal{F}_{i-1} \) be the information available up to the \( i \)th transaction at time \( t_{i-1} \). Denote the logarithm of the conditional durations by

\[
\phi_i = \ln(\psi_i) = \ln E(x_i|\mathcal{F}_{i-1})
\]

and assume

\[
x_i = e^{\phi_i} \varepsilon_i, \quad \varepsilon_i \sim \text{i.i.d., } E(\varepsilon_i) \equiv 1.
\]

The logACD model assumes that the log conditional mean evolves as

\[
\phi_i = \omega + \sum_{j=1}^{r} \alpha_j \phi_{i-j} + \sum_{k=1}^{s} \beta_k \varepsilon_{i-k} + \delta^T z^{(i-1)}
\]

with appropriate conditions for stationarity.

Note this form guarantees \( \psi_i > 0 \) even with included regression term \( \delta^T z^{(i-1)} \).

Engle (2000) “spreads, volume and quote arrivals are significant variables" [in earlier studies].
ACD model - Density of Excess Durations

Assume that the ‘excess durations’ $\varepsilon_i = x_i / \psi_i$ have a unit mean generalized Gamma distribution with density

$$q_{\varepsilon}(\varepsilon_i) = \frac{\theta}{\phi(\kappa, \theta) \Gamma(\kappa)} \left( \frac{\varepsilon_i}{\phi(\kappa, \theta)} \right)^{\kappa \theta - 1} \exp \left[ - \left( \frac{\varepsilon_i}{\phi(\kappa, \theta)} \right)^\theta \right]$$

and the conditional likelihood for $x_i$ given $\psi_i$ is

$$q(x_i | \psi_i) = \frac{\theta}{\phi(\kappa, \theta) \psi_i \Gamma(\kappa)} \left( \frac{x_i}{\phi(\kappa, \theta) \psi_i} \right)^{\kappa \theta - 1} \exp \left[ - \left( \frac{x_i}{\phi(\kappa, \theta) \psi_i} \right)^\theta \right]$$

Notes:

1. When $\theta = \kappa = 1$ get Exponential, when $\kappa = 1$ get Weibull and when $\theta = 1$ get Gamma.

2. Variety of hazard functions possible (increasing, unimodal, bathtub, decreasing) - see Bhatti.
IBM Data: Intraday Average Time Between Trades

Adjustment for time of day effects may not be the same across days.

William T.M. Dunsmuir (UNSW)
Modelling the Dynamics of Financial Transactions Data
August 8, 2009
Data Pre-Processing

Typically:

1. Remove any trades with time stamps outside trading hours.
2. Early trades in the day.
3. Adjust for time of day effects on average durations that are thought to be repeatable from day to day.
Commonly adjust for ‘predictable’ time of day effects:

\[ \hat{x}_i = x_i / f(t_i) \]

where \( f(t_i) \) is estimated halfhour by halfhour by averaging over data from all days or for each day of the week separately.

Example for IBM Data on following slide.
Note Friday appears different

Bauwens and Giot also note differences by day of week.

Alternative smoothing methods are linear splines or robust smoothers.

Results of Tang (2009) suggests that there are differences between days over the period - further work needed.
For example, Bauwens & Giot (2000) examined the way market-makers revise their beliefs about the prices by adding three variables in the autoregressive equation. These variables are related to characteristics of the trade process, which are the trading intensity, the average volume per trade and the spread when the past trades were made.
Apart from the question of intraday mean level adjustment being constant "the question of whether or not the estimated models are representative of daily market behavior remains open." (Bhatti, 2009). He concludes:

"Three of the stocks (GM, JNJ, SLB) provide support for homogeneity in interday trading dynamics with over two-thirds of the trading days having the same trading dynamics as the remainder of the sample—a surprising result if one believes that every trading day is different. On the other hand, the other three stocks (IBM, MCD, PG) provide some support for heterogeneity in interday trading dynamics with one-half to two-thirds of the trading days having different trading dynamics than the remainder of the sample."

Tang (2009) found that for IBM data the daily values of the ACD equation suggest autoregressive evolution from day to day but with low predictability.
Let \( y_i = P(t_i) - P(t_{i-1}) \) be a sequence of price changes in ticks and \( \mathcal{F}_{i-1} \) denote the information set available at the time \( t_i \) that transaction \( i \) takes place:

Models for the conditional distribution of the discrete price changes \( Y_i|\mathcal{F}_{i-1} \):

- **GARCH** (Engle 2000)
- **ACM** (Russell and Engle 2005)
- **Components of Price Change** (Activity, Direction and Size):
  - A, D, S components (Rydberg & Shephard 2003)
  - (AD), S components (Leisenfeld et al 2006)
GARCH Model for Price Change

Let the time adjusted returns be $r_i = y_i / x_i$ and

$$\sigma_i^2 = \text{Var}(r_i|\mathcal{F}_{i-1}, x_i)$$

Model for mean adjusted returns

$$r_i = \rho r_{i-1} + e_i + \phi e_{i-1}$$

Model for conditional variance

$$\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \psi_i^{-1} + \gamma_4 \xi_{i-1} + \gamma_5 G_{i-1} + \gamma_6 L_i$$
GARCH Model Terms

Note this model for price change volatility includes:

1. $\gamma_1$ for reciprocal duration ($x_i^{-1}$).
2. $\gamma_2$ for suprises in durations ($x_i / \psi_i$).
3. $\gamma_3$ for expected trade arrivals ($\psi_i$).
4. $\gamma_4$ for long run volatility $\zeta_i$, an EWMA of $r_i = y_i / x_i$.
5. $\gamma_5$ for previous bid-ask spread $G_{i-1}$.
6. $\gamma_6$ for $L_{i-1}$, an indicator of large volume ($>10000$) at time $t_{i-1}$.
Engle (2000) for 52,144 IBM stock transactions (November ’90 to January ’91):

1. Volatility contains both long and short term components.
2. Longer durations and longer expected durations are associated with lower volatilities (consistent with the Easley & O’Hara model).
3. Economic variables entered as expected by Easley & O’Hara model.
4. Higher bid-ask spreads and larger volumes both predict rising volatility.

Question: Is the GARCH model appropriate for highly discrete price changes?
Price Changes for IBM Example.

Histogram of IBM Price Changes less than 20 cents (40 days)
Liesenfeld et al 2006 observe:

"price jumps of more than ±5 ticks for 11% (JBX) and 6% (HAL) of the transactions, which supports our view that both modelling transaction returns as a continuous random variable and quantal response representation, are too crude to pick up the true nature of the dependent variable, and neglect valuable information about the true data generating process."

Two situations need to be considered:

1. Small price changes - ACM model
2. Moderate to large number of price change values - Component Models (e.g. ADS, (AD)S).
Assume that there are $k$ possible discrete values for the price changes. At the $i^{th}$ transaction, let $\pi_i$ denote the vector of probabilities for the $k-1$ nonzero price changes and let $\pi_0$ be the probability of zero price change. Let $h(\pi_i) = \log(\pi_i/(1-\pi_0))$ be the vector of log odds of a price change relative to no price change.

The Autoregressive Conditional Multinomial model of order $(p,q)$

$$h(\pi_i) = \sum_{j=1}^{p} A_j(x_{i-j} - \pi_{i-j}) + \sum B_j h(\pi_{i-j}) + Cz_i$$

where $A_j$ and $B_j$ denote the $j^{th}$ $(k-1) \times (k-1)$ parameter matrices, $z_i$ is an $r+1$ dimensional vector with 1 in the first element forming a constant and $r$ other explanatory variables, $C$ denotes a $(k-1) \times (r+1)$ conforming matrix of parameters.
Explanatory variables may contain predetermined variables such as characteristics of past trades including volume or spreads or information about the timing of trades.

Conclusions from Russell and Engle (2005):

"Both price returns and squared returns influence future durations and present and past durations affect price movements. The model exhibits reversals in transaction prices in the short run due to bid-ask bounce and clustering of large moves of either sign in the longer run. Evidence of symmetry in the dynamics of prices is presented, but the response to durations is clearly non-symmetric. It is found that the volatility per second of trades is highest for short duration trades and that expected returns are lower for longer duration trades."
ACM Model - Limits to Applicability

- Useful for stocks with small number of price changes. The Airgas example considered by Russell and Engle have only 0.07% price changes greater than 2 ticks.


  "A drawback of the ACM model is the necessity that all potential outcomes have to occur in the sample period to guarantee the identification and estimation of the true dimension of the multinomial process. This creates a serious limitation if the ACM is used for forecasting purposes ... the number of parameters increases with the outcome space."

- The ACM model will not be useful for data sets with more than a few ticks movement up or down. Even IBM which is highly traded has large range of up and down price changes making the ACM model unattractive and likely to be hard to identify (multivariate model with high dimensional state vector).
Components of Price Change (Activity, Direction and Size):

- A, D, S components (Rydberg & Shephard 2003)
- (AD), S components (Leisenfeld et al 2006)
Rydberg and Shepard (2003) proposed this as an alternative to ACM and GARCH. Let $Y_i$ be the price change between trades at time $t_{i-1}$ and $t_i$. Let

$$Y_i = A_i D_i S_i$$

where:

- **Activity**: $A_i = 1$ depending if a trade resulted in a price change and 0 otherwise,
- **Direction**: $D_i = -1$ if the price decreased and $D_i = 1$ if the price increased and 0 otherwise and
- **Size**: $S_i$ is the absolute amount of the price change.
Figure 3.2: Lend Lease Corporation stock price, September 2002.
Example Lend Lease

Price

Activity

Direction

Size
Model for Price Changes

\[ Pr(S_i = 0|\mathcal{F}_{i-1}) = Pr(A_i = 0|\mathcal{F}_{i-1}) \]

and for \( s_i \neq 0 \),

\[ Pr(S_i = s_i|\mathcal{F}_{i-1}) = Pr(A_i = 1|\mathcal{F}_{i-1}) \times \]
\[ \{ Pr(S_i = s_i|\mathcal{F}_{i-1}, A_i = 1, D_i = 1) Pr(D_i = 1|\mathcal{F}_{i-1}, A_i = 1) + \]
\[ Pr(S_i = -s_i|\mathcal{F}_{i-1}, A_i = 1, D_i = -1) Pr(D_i = -1|\mathcal{F}_{i-1}, A_i = 1) \} \]
Three Model Components for Price Changes

There are three pieces of modelling to be carried out:
1. $Pr(A_i|\mathcal{F}_{i-1})$ — the distribution of the activity process at time $t_i$ given the history to time $t_{i-1}$.
2. $Pr(D_i|\mathcal{F}_{i-1}, A_i = 1)$ — the distribution of the direction process at time $t_i$ given an Active trade, as well as the history to time $t_{i-1}$.
3. $Pr(S_i|\mathcal{F}_{i-1}, A_i = 1, D_i)$ — the distribution of the size process at time $t_i$ given an Active trade, its Direction, as well as the history to time $t_{i-1}$.
Model for the Binary Activity Process

\[ Pr(A_i | \mathcal{F}_{i-1}) = p(W_i) = \frac{\exp(W_i)}{1 + \exp(W_i)} \]

\[ W_i = z_i^T \beta + U_i \]

\[ U_i = \sum_{j=1}^{p} \phi_j U_{i-j} + \sigma e_{i-1} + \sigma \sum_{j=1}^{q} \theta_j e_{i-1-j} \]

\[ e_i = \frac{A_i - p(W_i)}{\sqrt{p(W_i)[1 - p(W_i)]}} \]
Exponential Family GLARMA models

Binomial, logit-link:

\[ I(\beta, \delta) = \sum_{i=1}^{n} \{ y_i W_i - n_i \ln(1 + \exp(W_i)) + c(y_i) \} \]

Negative binomial, log-link:

\[ I(\beta, \delta, \alpha) = \sum_{i=1}^{n} \left\{ y_i \ln\left( \frac{e^{W_i}}{\alpha + e^{W_i}} \right) - \alpha \ln\left( \frac{\alpha + e^{W_i}}{\alpha} \right) + c(y_i, \alpha) \right\} \]
Binomial Case, logit link \( p_i = \exp(W_i)/(1 + \exp(W_i)) \),

\[
e_i(\beta, \delta) = \frac{Y_i - n_i p_i}{\sqrt{n_i p_i (1 - p_i)}}
\]

Negative Binomial Case, log link, \( \mu_i = \exp(W_i) \)

\[
e_i(\beta, \delta, \alpha) = \frac{Y_i - \mu_i}{\sqrt{\mu_i + \mu_i^2 / \alpha}}
\]
Morris (2005) developed ADS models for three Australian companies Lend Lease Corporation Limited (LLC), Amcor Limited (AMC) and Coles Myer Limited (CML).

Some Conclusions:

1. Despite the three Australian companies chosen for the analysis coming from different industries, their stocks exhibited similar properties. The explanatory variables in the final models for the activity, direction and size processes were comparable for each stock.

2. However the lags with which the variables were statistically significant varied between components and between stocks.

3. The models selected typically represented economic theory very well. Three main effects were expected; increased likelihood of a large price change after a spell in trading, bid-ask bounce, and volatility clustering.
The ACD modelling of IBM transaction durations suggests possibility of day-to-day heterogeneity in model parameters. We also look at this question from the perspective of the price process - in the ADS component model.
Interday Variation Activity Series (IBM)
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>P-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>0.932</td>
<td>0.050</td>
<td>18.506</td>
<td>0.000</td>
</tr>
<tr>
<td>time difference</td>
<td>0.040</td>
<td>0.001</td>
<td>37.169</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt</td>
<td>-0.086</td>
<td>0.005</td>
<td>-18.473</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt1</td>
<td>-0.064</td>
<td>0.005</td>
<td>-13.729</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt2</td>
<td>0.024</td>
<td>0.005</td>
<td>5.172</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt3</td>
<td>0.015</td>
<td>0.005</td>
<td>3.205</td>
<td>0.001</td>
</tr>
<tr>
<td>Dt-1</td>
<td>-0.190</td>
<td>0.008</td>
<td>-24.933</td>
<td>0.000</td>
</tr>
<tr>
<td>Dt-2</td>
<td>-0.058</td>
<td>0.008</td>
<td>-7.475</td>
<td>0.000</td>
</tr>
<tr>
<td>Dt-3</td>
<td>-0.020</td>
<td>0.008</td>
<td>-2.655</td>
<td>0.008</td>
</tr>
<tr>
<td>Lt-1</td>
<td>-0.012</td>
<td>0.002</td>
<td>-5.067</td>
<td>0.000</td>
</tr>
<tr>
<td>Lt-2</td>
<td>0.008</td>
<td>0.002</td>
<td>3.081</td>
<td>0.002</td>
</tr>
<tr>
<td>Lt-3</td>
<td>0.012</td>
<td>0.003</td>
<td>4.792</td>
<td>0.000</td>
</tr>
<tr>
<td>theta1</td>
<td>0.177</td>
<td>0.006</td>
<td>28.688</td>
<td>0.000</td>
</tr>
<tr>
<td>theta2</td>
<td>0.075</td>
<td>0.006</td>
<td>12.121</td>
<td>0.000</td>
</tr>
<tr>
<td>theta3</td>
<td>0.062</td>
<td>0.006</td>
<td>9.901</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Interday Variation in Direction of Price Change (IBM)
### 40 Day Estimates Direction of Price Change (IBM)

<table>
<thead>
<tr>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>P-val</th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>P-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>-0.226</td>
<td>0.044</td>
<td>-5.087</td>
<td>0.000</td>
<td>LDt-1</td>
<td>-0.039</td>
<td>0.003</td>
</tr>
<tr>
<td>time differ</td>
<td>-0.007</td>
<td>0.001</td>
<td>-7.380</td>
<td>0.000</td>
<td>LDt-2</td>
<td>-0.023</td>
<td>0.003</td>
</tr>
<tr>
<td>ln_vol1</td>
<td>0.039</td>
<td>0.006</td>
<td>6.841</td>
<td>0.000</td>
<td>LDt-3</td>
<td>-0.012</td>
<td>0.003</td>
</tr>
<tr>
<td>ln_vol2</td>
<td>-0.007</td>
<td>0.006</td>
<td>-1.064</td>
<td>0.287</td>
<td>DT1</td>
<td>0.566</td>
<td>0.085</td>
</tr>
<tr>
<td>ln_vol3</td>
<td>0.014</td>
<td>0.006</td>
<td>2.320</td>
<td>0.020</td>
<td>DT2</td>
<td>0.410</td>
<td>0.053</td>
</tr>
<tr>
<td>At-1</td>
<td>0.030</td>
<td>0.016</td>
<td>1.901</td>
<td>0.057</td>
<td>DT3</td>
<td>-0.115</td>
<td>0.028</td>
</tr>
<tr>
<td>At-2</td>
<td>-0.013</td>
<td>0.016</td>
<td>-0.801</td>
<td>0.423</td>
<td>theta1</td>
<td>-0.852</td>
<td>0.084</td>
</tr>
<tr>
<td>At-3</td>
<td>-0.045</td>
<td>0.016</td>
<td>-2.864</td>
<td>0.004</td>
<td>theta2</td>
<td>-0.091</td>
<td>0.062</td>
</tr>
<tr>
<td>Lt-1</td>
<td>-0.005</td>
<td>0.003</td>
<td>-1.668</td>
<td>0.095</td>
<td>theta3</td>
<td>0.176</td>
<td>0.027</td>
</tr>
<tr>
<td>Lt-2</td>
<td>0.000</td>
<td>0.003</td>
<td>-0.038</td>
<td>0.969</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lt-3</td>
<td>0.004</td>
<td>0.003</td>
<td>1.461</td>
<td>0.144</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Interday Variation in Size of Price Change (IBM)
## 40 Day Estimates Size of Price Change (IBM)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>Z</th>
<th>P-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>-1.433</td>
<td>0.048</td>
<td>-29.920</td>
<td>0.000</td>
</tr>
<tr>
<td>time difference</td>
<td>0.024</td>
<td>0.001</td>
<td>31.570</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt</td>
<td>0.185</td>
<td>0.004</td>
<td>51.788</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt1</td>
<td>0.036</td>
<td>0.004</td>
<td>9.744</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt2</td>
<td>0.032</td>
<td>0.004</td>
<td>8.598</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_volt3</td>
<td>0.024</td>
<td>0.004</td>
<td>6.583</td>
<td>0.000</td>
</tr>
<tr>
<td>At-1</td>
<td>0.054</td>
<td>0.010</td>
<td>5.552</td>
<td>0.000</td>
</tr>
<tr>
<td>At-2</td>
<td>0.040</td>
<td>0.010</td>
<td>4.200</td>
<td>0.000</td>
</tr>
<tr>
<td>At-3</td>
<td>0.023</td>
<td>0.010</td>
<td>2.344</td>
<td>0.019</td>
</tr>
<tr>
<td>Dt-1</td>
<td>0.125</td>
<td>0.006</td>
<td>21.782</td>
<td>0.000</td>
</tr>
<tr>
<td>Dt-2</td>
<td>0.076</td>
<td>0.006</td>
<td>12.728</td>
<td>0.000</td>
</tr>
<tr>
<td>Dt-3</td>
<td>0.050</td>
<td>0.006</td>
<td>8.542</td>
<td>0.000</td>
</tr>
<tr>
<td>theta1</td>
<td>0.237</td>
<td>0.005</td>
<td>49.787</td>
<td>0.000</td>
</tr>
<tr>
<td>theta2</td>
<td>0.167</td>
<td>0.005</td>
<td>34.934</td>
<td>0.000</td>
</tr>
<tr>
<td>theta3</td>
<td>0.105</td>
<td>0.005</td>
<td>22.253</td>
<td>0.000</td>
</tr>
<tr>
<td>alpha</td>
<td>0.776</td>
<td>0.007</td>
<td>116.232</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Conclusions: Interday Stability of ADS Model Parameters for IBM Price Changes

1. Overwhelming evidence that model parameters are not constant from day to day - LRT’s are large and highly significant rejecting the null hypothesis of constant parameters from day to day.

2. For the activity process $A_t$ time difference ($x_t$) and the excess price change at two lags show significant trends over the 40 days.

3. For the direction of price change process $D_t$ the intercept, and ‘large direction’ indicating if the previous price changes moved by more than one tick show significant trends.

4. For the size of price change process $S_t$ time difference and lagged $A_{t-1}$ and $D_{t-2}$ and $D_{t-3}$ show significant trends.

5. Overall, many parameters are not varying from day to day but the ones that are cannot be ignored.
Decompose the overall process of transaction price changes into three components.

The first component determines the direction of the process (positive price change, negative price change, or no price change) and will be specified as a dynamic multinomial response model (an ACM model).

Given the direction of the price change, count data processes determine the size of positive and negative price changes, representing the second and third component of the model.
Models \((AD)_i = A_i D_i\) in a single process of multinomial outcomes taking values \(-1, 0, 1\) if the price decreased, stayed the same or increased. Direction of price changes has trinomial distribution for

\[
P(Y_i < 0|\mathcal{F}_{i-1}), \quad P(Y_i = 0|\mathcal{F}_{i-1}) \quad \text{and} \quad P(Y_i > 0|\mathcal{F}_{i-1}).
\]

Size of the price changes conditional on the price direction, are defined by two processes

\[
P(Y_i = y_i|Y_i < 0, \mathcal{F}_{i-1}) \quad \text{and} \quad P(Y_i = y_i|Y_i > 0, \mathcal{F}_{i-1})
\]
both of these being on the strictly positive integers.
They use a GLARMA model based on a truncated-at zero Negative Binomial distribution.
Generalized Autoregressive Score model

GAS model of Creal et al (2008) provides a unifying model for single source of error (observation driven) processes which:

- Includes many existing models such as GARCH, ACD, ACI, ADS, GLARMA
- Allows a new range of observation driven models.

Note:

1. Observation driven models have a state equation relying on a single source of error using past observations.
2. Parameter Driven models allow for an additional unobserved source or random variation in the state equation.
3. All the models considered in this talk are observation driven.
Let $Y_t$ be the dependent variable (possible vector) at time $t$, $f_t$ be the time varying parameter vector, $x_t$ be the vector of covariates (exogenous). The GAS model has two components

Observations Density: $p(y_t | Y_1^{t-1}, F_1^{t-1}, X_1^{t-1}; \theta)$

$$f_t = \omega + \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=1}^{q} B_i f_{t-j}$$

Parameters:

$$\omega = \omega(\theta), \ A_i = A_i(\theta), \ B_i = B_i(\theta)$$
Choice of driving mechanism

The driving mechanism $s_t$ in equation (2) is based on scaling the score

$$s_t = S_{t-1} \nabla_t, \quad \nabla_t = \frac{\partial \ln p(y_t | \cdots)}{\partial f_{t-1}}$$

where $S_{t-1} = S(t, Y_{1}^{t-1}, X_1^t, F_1^{t-1}; \theta)$ can be (and has been) chosen in several ways:

1. Identity: $S_{t-1} = I$
2. Inverse of information matrix

$$S_{t-1} = E_{t-1}(\nabla_t \nabla_t') = \left[ -E \left( \frac{\partial^2 \ln p(y_t | \cdots)}{\partial f_{t-1} \partial f_{t-1}'} \right) \right]^{-1}$$
We now have a range of models for the durations between transactions and the price change process associated with high frequency financial data.

All these models rely on non-Gaussian, often integer valued outcome distributions together with an ARMA type dynamic relationship for the state equation driving the outcome distribution.

There are now numerous applications of these models to a diversity of financial series - a few of these have been discussed in this talk.

However, a review of all such studies and their findings, particularly in regard to their conclusions concerning market infrastructure would be useful (cf metaanalysis in public health issues).

There remain some substantial modelling and theory issues.
1. Most applications assume that the data adjustments (e.g. diurnal effects), model structure and model parameters are the same for all days. However there is evidence to suggest this is not always correct.

2. Further empirical research is required concerning this on much longer and more diverse markets and series. Methods of Dunsmuir (2009) could be extended to test that daily models can be simplified.

3. Hierarchical models in which ACD/ACM/ADS/GAS model parameters are themselves modelled by day to day time series models should be developed and evaluated. Bayesian models and methods may be useful for these.

4. Forecasting accuracy and utility needs to be assessed more widely.

5. Compare ADS and Hurdle Models.
Transactions data is voluminous so asymptotic theory for estimation and testing should be useful. However very little is known about:

- The ergodicity/stability of the dynamic models for discrete valued series (exceptions include simple models for Poisson in Davis et al 2003, 2005 and Fokianos et al 2009 and binary case Streett (2000)).
- Consistency and asymptotic normality of likelihood estimates.

Computational: Fitting GLARMA type dynamic structures is now workable for long series of the type encountered here. But speed needs to be improved for simulation studies and larger data sets.

Forecasting methods for discrete series for several steps ahead is computationally intensive - further development needed.
I would like to thank Professor Shiba for inviting me to make this presentation and to the organizers for their support to participate in the 3rd CFEE Conference on Financial Engineering Education.

Chad Bhatti kindly provided his IBM data series and his R/Fortran code for fitting the generalized gamma distribution model.

Students at UNSW have contributed over the past 4 years to material related to this talk: Bo Wang, Daniel Drescher, Yingfang Tang, Haitao Shao.


