Modelling the Dynamics of Financial Transactions Data

William T.M. Dunsmuir

UNSW

August 8 2009

Talk Objective and Outline

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

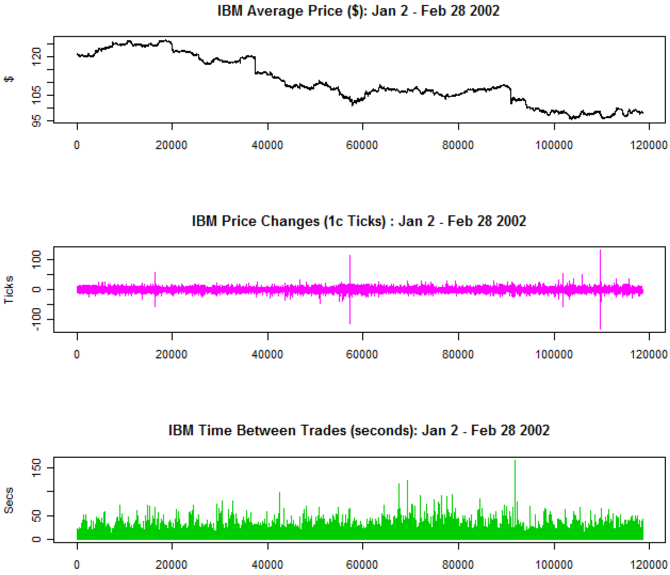
Outline:

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

"So do flux and reflux – the rythm of change – alternate and persist in everything under the sky."

- T.Hardy 'Tess of the D'Urbervilles'

Example IBM Jan 1 to Feb 28, 2002



Ultra High Frequency Data

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 - Price Up or Down?.
- Size of price change.

Model for Durations and Price Changes

Want an econometric model for the joint process of transaction arrivals and tick-by-tick price movements.
Let:

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

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

$$f(y_i, x_i | y^{(i-1)}, x^{(i-1)}, z^{(i)}) = g(y_i | y^{(i-1)}, x^{(i)}, z^{(i)}) \times q(x_i | y^{(i-1)}, x^{(i-1)}, z^{(i)})$$

Use of Models:

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.as well as previous price changes and times between trades.

Models for Times Between Trades

Two types of models proposed for the duration between successive trades:

1. ACD models (Engel)
2. Discrete time BIN model (Rydberg and Shephard)

ACD Type Models for Durations

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 ith 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

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

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

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

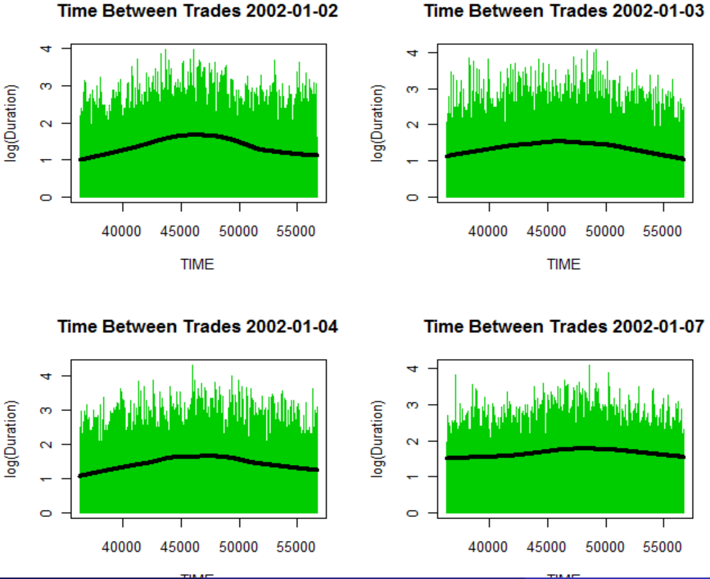
and the conditional likelihood for x_i given ψ_i is

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

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



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.

IBM Example: Adjusment for Time of Day Effects

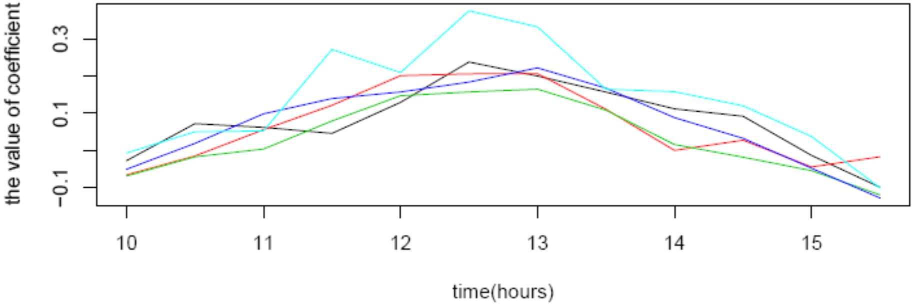
Commonly adjust for ‘predictable’ time of day effects:

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

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.

Intraday Average Time Between Durations



Time of Day Adjustments

- 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.

Examples of ACD empirical applications.

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:

1. trading intensity
2. average volume per trade
3. spread when the past trades were made.

Interday Stability of ACD models.

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.

Models for Price Changes

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

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

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. γ_1 for reciprocal duration (x_i^{-1}).
2. γ_2 for suprises in durations (x_i/ψ_i).
3. γ_3 for expected trade arrivals (ψ_i).
4. γ_4 for long run volatility ξ_i, an EWMA of $r_i = y_i/x_i$.
5. γ_5 for previous bid-ask spread G_{i-1}.
6. γ_6 for L_{i-1}, an indicator of large volume ($>$10000) at time t_{i-1}.

GARCH Model - Some Conclusions

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 Easly & 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.

Need for Discrete Valued Models

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).

ACM Model for Price Changes - Definitions

Assume that there are k possible discrete values for the price changes.

At the i^{th} transaction let:

- π_i denote the vector of probabilities for the $k-1$ nonzero price changes
- π_0 be the probability of zero price change
- the vector of log odds of a price change relative to no price change

$$h(\pi_i) = \log(\pi_i / (1 - \pi_0))$$

ACM Model for Price Changes - Model Form

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.

ACM Model - some results.

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.
- Liesenfeld et al (2006):

"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).

Models for Components of Price Changes

- Components of Price Change (Activity, Direction and Size):
  - A, D, S components (Rydberg & Shephard 2003)
  - (AD), S components (Leisenfeld et al 2006)

Description of ADS model

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 $D_i = 1$ is the price increased and 0 otherwise and
- Size: S_i is the abosolute amount of the price change.

Example Lend Lease

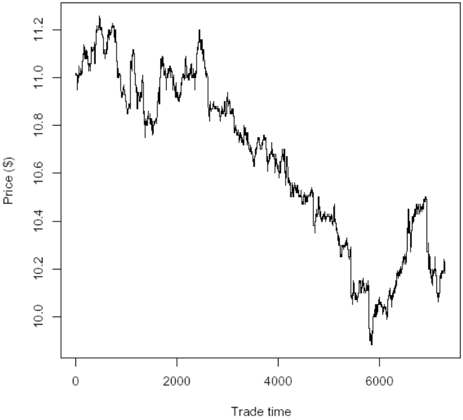
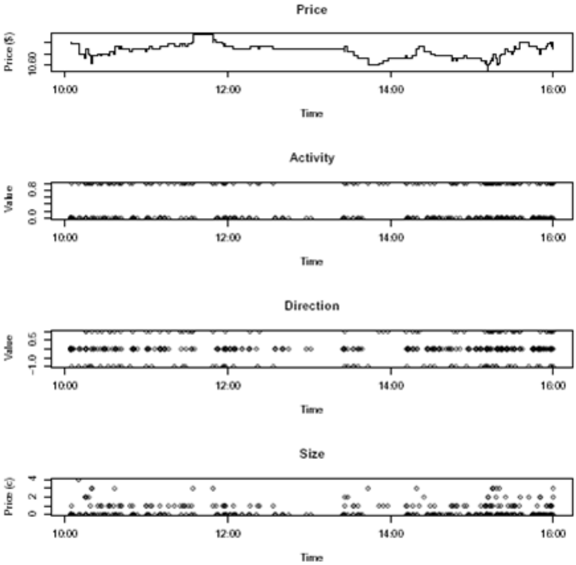


Figure 3.2: Lend Lease Corporation stock price, September 2002.

Example Lend Lease



Model for Price Changes

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

and for $s_i \neq 0$,

$$\begin{aligned} Pr(S_i = s_i|\mathcal{F}_{i-1}) = Pr(A_i = 1|\mathcal{F}_{i-1}) \times \\ \{Pr(S_i = s_i|\mathcal{F}_{i-1}, A_i = 1, D_i = 1) Pr(D_i = 1|\mathcal{F}_{i-1}, A_i = 1) + \\ Pr(S_i = -s_i|\mathcal{F}_{i-1}, A_i = 1, D_i = -1) Pr(D_i = -1|\mathcal{F}_{i-1}, A_i = 1)\} \end{aligned}$$

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

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

$$W_i = z_i^{\mathrm{T}}\beta + U_i$$

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

$$e_i = \frac{A_i - p(W_i)}{\sqrt{p(W_i)[1 - p(W_i)]}}$$

Exponential Family GLARMA models

Binomial, logit-link:

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

Negative binomial, log-link:

$$l(\beta, \delta, \alpha) = \sum_{i=1}^{n} \left\{ y_i \ln(\frac{e^{W_i}}{\alpha + e^{W_i}}) - \alpha \ln(\frac{\alpha + e^{W_i}}{\alpha}) + c(y_i, \alpha) \right\}$$

Exponential Family GLARMA models

- 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, $\mu_i = \exp(W_i)$

$$e_i(\beta,\delta,\alpha) = \frac{Y_i - \mu_i}{\sqrt{\mu_i + \mu_i^2/\alpha}}$$

Computational Details - 1

Let $\mathcal{F}_t = (Y^{(t-1)}, X^{(t)})$ be the past of an observed count process and the past and present of the regressor variables and assume that the conditional distribution of $Y_t|\mathcal{F}_t$ is denoted by $f(y_t|W_t;\delta)$ in terms of some state process W_t, defined in terms of a vector of parameter δ and elements from $\mathcal{F}_t$.

We are considering regression models of the type

$$W_t = x_t^T \beta + Z_t$$

$$Z_t = \sum_{i=1}^{\infty} \gamma_i e_{t-i}$$

where the infinite moving average in this model can be specified in terms of a finite number of parameters in many ways.

Computational Details - 2

We consider distributed lag structures that are generated by the linear predictor for autoregressive-moving average processes of the form

$$\sum_{i=1}^{\infty} \gamma_i u^i = (1 - \sum_{i=1}^{p} \phi_i u^i)^{-1}(1 + \sum_{i=1}^{q} \theta_i u^i) - 1.$$

Z_t is computed using the autoregressive moving average recursions

$$Z_t = \phi_1(Z_{t-1} + e_{t-1}) + \cdots + \phi_p(Z_{t-p} + e_{t-p}) + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q},$$

and throughout, for $t \leq 0$, set $e_t = 0$, $Z_t = 0$ to initialize the recursions.

Computational Details - 3

For the Poisson response distribution and log link

$$\log(\mu_t) = W_t(\delta),$$

the log-likelihood is

$$l(\delta) = \sum_{t=1}^{n} \{ Y_t W_t(\delta) - e^{W_t(\delta)} - \log(y_t!) \}.$$

The first and second derivatives of l are given by the following expressions

$$\frac{\partial l}{\partial \delta} = \sum_{t=1}^{n} (Y_t - \mu_t) \frac{\partial W_t}{\partial \delta}$$

$$\frac{\partial^2 l}{\partial \delta \partial \delta^T} = \sum_{t=1}^{n} \left[(Y_t - \mu_t) \frac{\partial^2 W_t}{\partial \delta \partial \delta^T} - \mu_t \frac{\partial W_t}{\partial \delta} \frac{\partial W_t}{\partial \delta^T} \right]$$

Computational Details - 4

The core part of our software is the recursive calculation of $\frac{\partial W_t}{\partial \delta}$ and $\frac{\partial^2 W_t}{\partial \delta \partial \delta^T}$:
Example:

$$\frac{\partial W_t}{\partial \delta} = \frac{\beta}{\partial \delta} + \frac{\partial Z_t}{\partial \delta},$$

where

$$\frac{\partial Z_t}{\partial \delta} = \sum_{i=1}^{p} \frac{\partial \phi_i}{\partial \delta}(Z_{t-i} + e_{t-i}) + \sum_{i=1}^{p} \phi_i \left(\frac{\partial Z_{t-i}}{\partial \delta} + \frac{\partial e_{t-i}}{\partial \delta} \right) + \sum_{i=1}^{q} \frac{\partial \theta_i}{\partial \delta} e_{t-i} + \sum_{i=1}^{q} \theta_i \frac{\partial e_{t-i}}{\partial \delta}.$$

Computational Details - 5

- Maximisation of Likelihood based on Approximate Fisher Scoring or Newton Raphson
- Standard errors using $-\left(\frac{\partial^2 l(\hat{\delta})}{\partial\delta\partial\delta^T}\right)^{-1}$
- Approximate normal distribution and large sample χ^2 distribution for LRT proved only in very simple cases (constant regression, MA-1).

Computational Details - 6

Current software covers:

- Poisson, Negative Binomial, Binary and Binomial responses - modification for other response distributions is relatively straightforward.
- Pearson and Score type (see later) innovations for recursive equation for Z_t, and hence W_t, are available.
- Recent modification to allow news impact (asymmetric response to past innovations in the state equation) and to allow trinomial responses require modifications to the basic recursive 'engine').

Empirical Application of Price Change Model to 3 Australian Stocks

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.

Interday Stability of Components of Price Change Models.

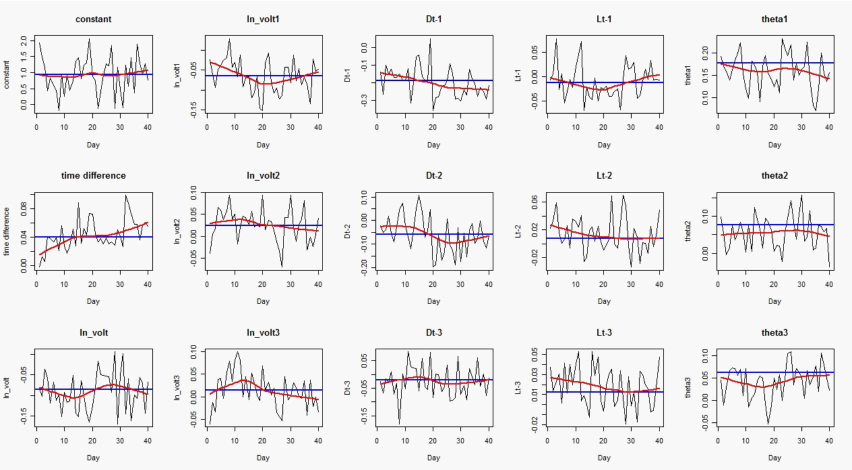
The ACD modelling of IBM transaction durations suggests possibility of day-to-day heterogeneity in model parameters.

We also look at this question from the perspective of the price process - in the ADS component model.

40 Day Estimates Activity of Price Change (IBM)

	Estimate	SE	Z	P-val
constant	0.932	0.050	18.506	0.000
time difference	0.040	0.001	37.169	0.000
ln_volt	-0.086	0.005	-18.473	0.000
ln_volt1	-0.064	0.005	-13.729	0.000
ln_volt2	0.024	0.005	5.172	0.000
ln_volt3	0.015	0.005	3.205	0.001
Dt-1	-0.190	0.008	-24.933	0.000
Dt-2	-0.058	0.008	-7.475	0.000
Dt-3	-0.020	0.008	-2.655	0.008
Lt-1	-0.012	0.002	-5.067	0.000
Lt-2	0.008	0.002	3.081	0.002
Lt-3	0.012	0.003	4.792	0.000
theta1	0.177	0.006	28.688	0.000
theta2	0.075	0.006	12.121	0.000
theta3	0.062	0.006	9.901	0.000

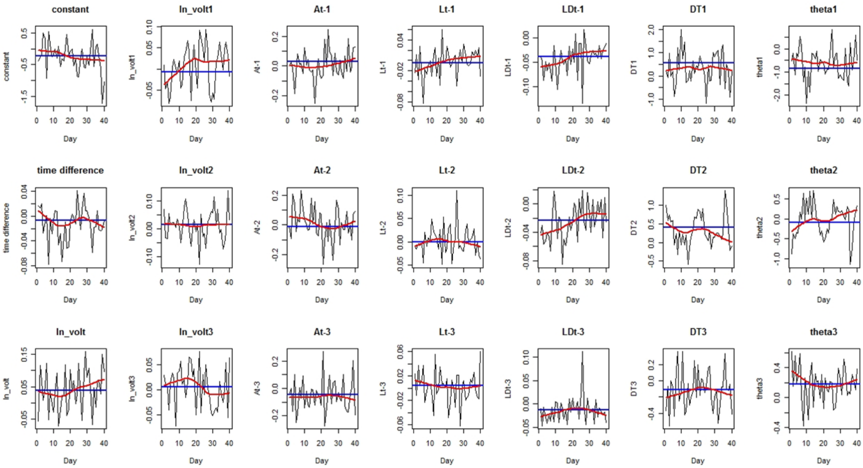
Interday Variation Activity Series (IBM)



40 Day Estimates Direction of Price Change (IBM)

	Estimate	SE	Z	P-val			Estimate	SE	Z	P-val
constant	-0.226	0.044	-5.087	0.000		LDt-1	-0.039	0.003	-13.140	0.000
time diffe	-0.007	0.001	-7.380	0.000		LDt-2	-0.023	0.003	-8.318	0.000
ln_volt	0.039	0.006	6.841	0.000		LDt-3	-0.012	0.003	-4.585	0.000
ln_volt1	-0.007	0.006	-1.064	0.287		DT1	0.566	0.085	6.626	0.000
ln_volt2	0.014	0.006	2.320	0.020		DT2	0.410	0.053	7.797	0.000
ln_volt3	0.005	0.006	0.950	0.342		DT3	-0.115	0.028	-4.134	0.000
At-1	0.030	0.016	1.901	0.057		theta1	-0.852	0.084	-10.088	0.000
At-2	-0.013	0.016	-0.801	0.423		theta2	-0.091	0.062	-1.479	0.139
At-3	-0.045	0.016	-2.864	0.004		theta3	0.176	0.027	6.462	0.000
Lt-1	-0.005	0.003	-1.668	0.095						
Lt-2	0.000	0.003	-0.038	0.969						
Lt-3	0.004	0.003	1.461	0.144						

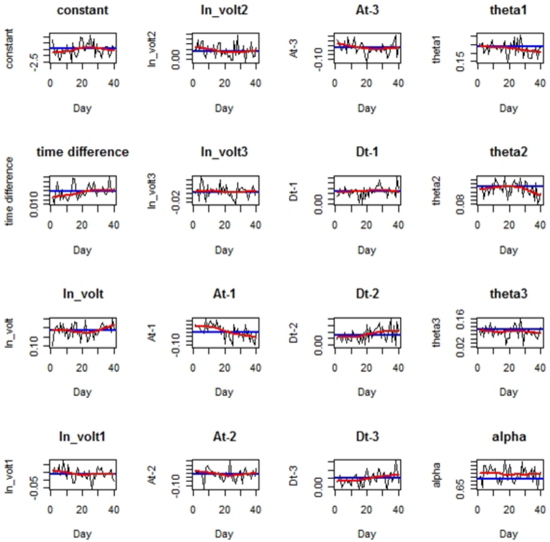
Interday Variation in Direction of Price Change (IBM)



40 Day Estimates Size of Price Change (IBM)

	Estimate	SE	Z	P-val
constant	-1.433	0.048	-29.920	0.000
time difference	0.024	0.001	31.570	0.000
ln_volt	0.185	0.004	51.788	0.000
ln_volt1	0.036	0.004	9.744	0.000
ln_volt2	0.032	0.004	8.598	0.000
ln_volt3	0.024	0.004	6.583	0.000
At-1	0.054	0.010	5.552	0.000
At-2	0.040	0.010	4.200	0.000
At-3	0.023	0.010	2.344	0.019
Dt-1	0.125	0.006	21.782	0.000
Dt-2	0.076	0.006	12.728	0.000
Dt-3	0.050	0.006	8.542	0.000
theta1	0.237	0.005	49.787	0.000
theta2	0.167	0.005	34.934	0.000
theta3	0.105	0.005	22.253	0.000
alpha	0.776	0.007	116.232	0.000

Interday Variation in Size of Price Change (IBM)



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.

Hierarchical Model Formulation?

Level 1 Observations from exponential family

Level 2 On day d of trading the state equation is

$$W_{d,t} = x_{d,t}^T \beta_d + Z_{d,t}$$
$$Z_{d,t} = \sum_{i=1}^{\infty} \gamma_i(\tau_d) e_{d,t-i}$$

Level 3 Let $\delta_d = (\beta_d, \tau_d)$ be the parameters for day d and assume that δ_d follows a multiple time series model with exogenous variables.

Hurdle Model of Liesenfeld, Nolte and Pohlmeier (2006)

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.

Hurdle Model - Details

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

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

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

$$P(Y_i = y_i|Y_i < 0, \mathcal{F}_{i-1}) \text{ and } P(Y_i = y_i|Y_i > 0, \mathcal{F}_{i-1})$$

both of these being on the strictly positive integers.

PDF for Hurdle Model

$$P(Y_i = y_i|\mathcal{F}_{i-1}) = \begin{cases} P(Y_i = y_i|Y_i < 0, \mathcal{F}_{i-1})P(Y_i < 0|\mathcal{F}_{i-1}) & \text{if } y_i < 0 \\ P(Y_i = 0|\mathcal{F}_{i-1}) & \text{if } y_i = 0 \\ P(Y_i = y_i|Y_i > 0, \mathcal{F}_{i-1})P(Y_i > 0|\mathcal{F}_{i-1}) & \text{if } y_i > 0 \end{cases}$$

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.

Details of GAS model

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

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

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

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

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

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. Indentity: $S_{t-1} = I$
2. (Score type) Inverse of information matrix

$$S_{t-1} = \left[-E\left(\frac{\partial^2 \ln p(y_t | \cdots)}{\partial f_{t-1} \partial f'_{t-1}} \right) \right]^{-1}$$

3. Square root of information matrix.

Score versus Pearson Residuals - Form

	Score (option 2)	Pearson	Parameterized
Poisson	$\frac{Y_t - \mu_t}{\mu_t}$	$\frac{Y_t - \mu_t}{\mu_t^{0.5}}$	$\frac{Y_t - \mu_t}{\mu_t^{\lambda}}$
Negative Binomial	$\frac{Y_t - \mu_t}{\mu_t}$	$\frac{Y_t - \mu_t}{\sqrt{\mu_t + \mu_t^2/\alpha}}$	-
Binomial	$\frac{Y_t - n_t p_t}{n_t p_t(1 - p_t)}$	$\frac{Y_t - n_t p_t}{\sqrt{n_t p_t(1 - p_t)}}$	-

Score versus Pearson Residuals -Impact

Simulation of simple linear trend and Poisson counts with lag 1 AR or lag 1 MA suggests:

1. When Pearson residuals true and Pearson or Score residuals assumed regression parameters have similar biases and standard errors but the autocorrelation parameter estimated assuming Score residuals is biased away from zero.
2. When Score residuals true the autocorrelation parameter estimated assuming Pearson residuals is biased towards zero.
3. The standard errors using Pearson residuals in the fitted model are less biased than the Score residuals regardless of how the residuals were generated - reason not understood.

Score versus Pearson Residuals -Asymptotics

- In his original formulation Shephard (1995) suggested using Score type residuals in the Poisson case.
- Pearson residuals may be more appealing being zero mean, unit variance, martingale differences.
- Asymptotics for constant mean, lag 1 Moving average available ('and that's all folks'!).
- The process $\{W_t\}$ is uniformly ergodic for the case $\lambda = 1$. Hence, there exist unique stationary distributions for both the log-mean and mean processes in this case of Score residuals.
- For $1/2 \leq \lambda < 1$, there exists a stationary distribution, yet the uniqueness of such a distribution is currently unknown.
- For other values, $0 < \lambda < 1/2$, the stability properties of the process are not yet understood.
- For $\lambda = 0$ process diverges.

Concluding Summary

1. We now have a range of models for the durations between transactions and the price change process associated with high frequency financial data.
2. 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.
3. There are now numerous applications of these models to a diversity of financial series - a few of these have been discussed in this talk.
4. 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).
5. There remain some subtantial modelling and theory issues.

Open Research Questions - Modelling

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.

Open Research Questions - Theory

1. 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.
2. 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.
3. Forecasting methods for discrete series for several steps ahead is computationally intensive – Rydberg and Shephard (2003) discuss this. Further development is needed. Recent work of Martin and McCabe relevant here.

Acknowledgements

- This talk, in an earlier form, was presented at the 3rd CFEE Conference on Financial Engineering Education at Hitotsubashi University in Tokyo in August 2009.
- 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, Brad Morris, Daniel Drescher, Yingfang Tang, Haitao Shao, Qing Zou and Lei Ding.

References - 1

- Allen, D., F. Chan, M. McAleer & S. Peiris (2008). Finite sample properties of the qmle for the log-ACD model: application to australian stocks, Journal of Econometrics 147 (1), 163–183.
- Bauwens, L. & P. Giot, The logarithmic ACD model: an application to the bid-ask quote process of three nyse stocks, Annales D' Économie Et De Statistique 60 (2000) 117–149.
- Bhatti, C.R. (2009): On the interday homogeneity in the intraday rate of trading. Mathematics and Computers in Simulation, 79, 2250-2257.Davis, R.A, W.T.M. Dunsmuir and S.B. Streett (2003) "Observation driven models for Poisson counts", Biometrika, 90, 777-790.
- Cipollini, F., R. F. Engle, & G. M. Gallo (2006). Vector multiplicative error models: representation and inference. NBER Working Paper.

References - 2

- Creal, D., Siem Jan Koopman, André Lucas (2008) "A General Framework for Observation Driven Time-Varying Parameter Models" TI 2008-108/4 Tinbergen Institute Discussion Paper.
- Davis, R.A, W.T.M. Dunsmuir and S.B. Streett (2003) “Observation driven models for Poisson counts”, Biometrika, 90, 777-790.
- Davis, R.A, W.T.M. Dunsmuir and S.B. Streett (2005) “Maximum likelihood estimation for an observation driven model for Poisson counts”, Methodology, Computing and Applied Probability, 7, 149-159.
- Dunsmuir, W.T.M., J. Leung and X. Liu (2004) “Extensions of Observation Driven Models for Time Series of Counts”, Proceedings of the International Sri Lankan Statistical Conference: Visions of Futuristic Methodologies, eds Basil M. de Silva and Nitis Mukhopadhyay, RMIT University and University of Peradeniya

References - 3

- Dunsmuir, W.T.M. (2009) "Inference and Computational Methods for Regression Modelling in Multiple Time Series of Counts with Random Effects", presented at the International Symposium in Statistics on Inferences in Generalized Linear Longitudinal Mixed Models, St John's, Canada. In preparation for journal submission.
- Easley D., O'Hara M. (1992): Time and the Process of Security Price Adjustment. The Journal of Finance 19, pp.69-90.
- Engle R (2000) The econometrics of ultra-high-frequency data. Econometrica 68(1):1–22.
- Fokianos, C., A. Rahbek & D. Tjostheim (2009) "Poisson autoregression", draft manuscript.
- Liesenfeld, R., Nolte, I., Pohlmeier, W. (2006) "Modelling financial transaction price movements: a dynamic integer count data model, Empirical Economics 30:795–825
- Morris, B. L. (2005) "Application and comparison of models for high frequency price data", Honours Thesis, UNSW School of Mathematics and Statistics.

- Russell, J. R. and R. F. Engle (2005). A discrete-state continuous-time model of financial transactions prices and times: the autoregressive conditional multinomial-autoregressive conditional duration model. Journal of Business & Economic Statistics 23 (2), 166-180.
- Rydberg, T. H. and N. Shephard (2003). Dynamics of trade-by-trade price movements: decomposition andmodels. Journal of Financial Econometrics 1, 2-25.
- Streett, S. (2000) "Some Observation Driven Time Series", Unpublished PhD dissertation, Department of Statistics, Colorado State University.
- Tang, Y. (2009) "Estimating Log-ACD Models for Stock Transaction Data", Master of Statistics Report, UNSW School of Mathematics and Statistics.
- Wang, B. (2004) "GLARMA Models and Stock Price Dynamics", Master of Statistics Report, UNSW School of Mathematics and Statistics.