The Econometrics of High Frequency Data

Jeffrey R. Russell

Prepared for the High Frequency Data Tutorial

Organized by

Mathematical Sciences, National University of Singapore
and
School of Economics and Social Sciences, Singapore Management University

Tentative Plan

- Day 1 Introduction to high frequency data
  - Data features
  - Market microstructure.
  - Fixed Interval Analysis
- Day 2 A point process approach – the ACD model
- Day 3 Marked point process modeling.
  - General approach
  - Volatility Models
  - Discrete price movements – the ACM-ACD model
- Day 4 Measuring and modeling transaction cost (slides forthcoming)
  - Existing measures
  - Full information transaction cost
  - Decomposing the spread with unobserved components.

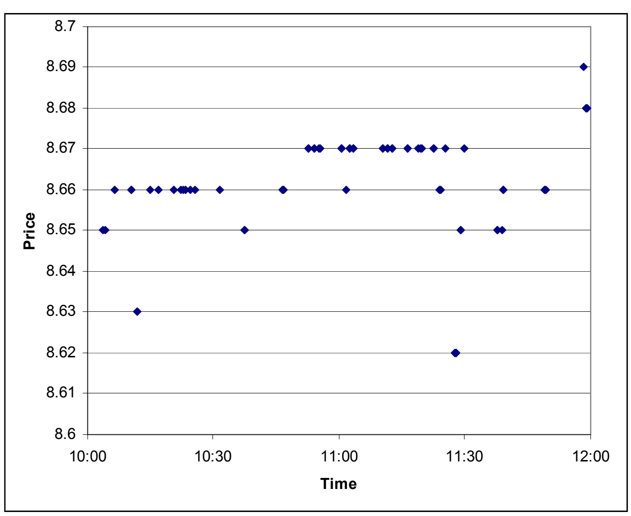
2

Introduction

- Characteristics of High Frequency Data
  - Irregularly Spaced
  - Diurnal Patterns
  - Separate prices for buying and selling
  - Discrete prices
  - Highly dependent
- We will use the Airgas stock traded on the NYSE throughout the week.

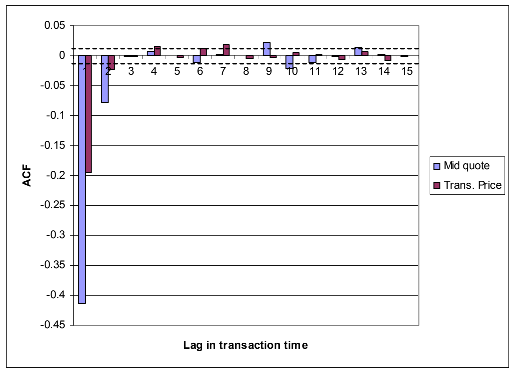
3

2 hours of transaction price data



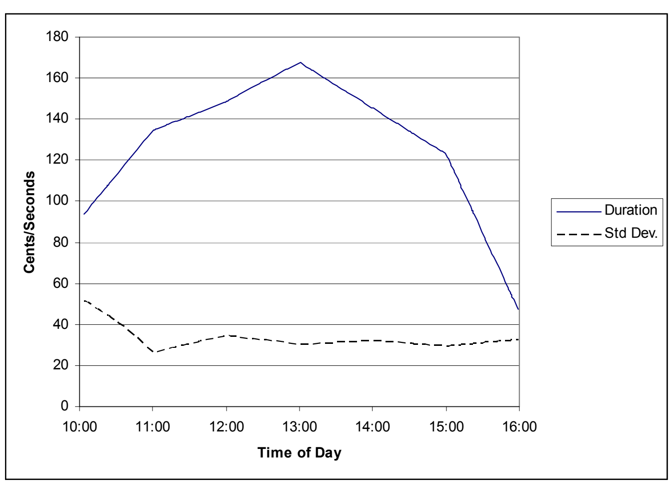
4

Autocorrelation in prices



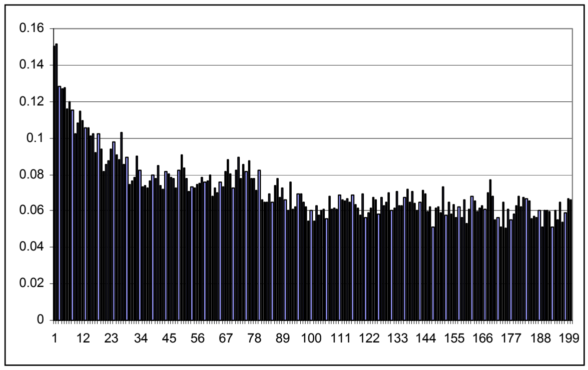
5

Diurnal Patterns



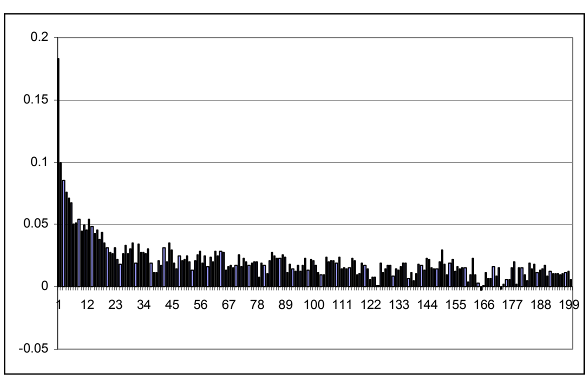
6

Autocorrelations in durations



7

Autocorrelations in (log) volume



8

Economic Data

- Quotes
- Transaction prices
- Volume
- Time stamps
- Orders

9

Economic Questions

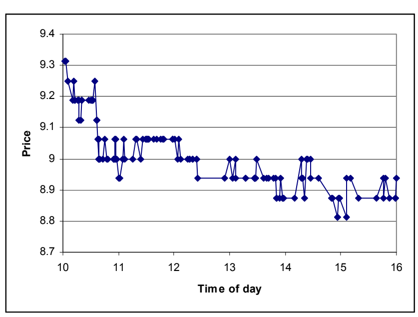
- Liquidity/Quality of Execution
- Forecasting Order Flow
- Hedging/Pricing
- Price Discovery
- Volatility/Risk
- Correlation

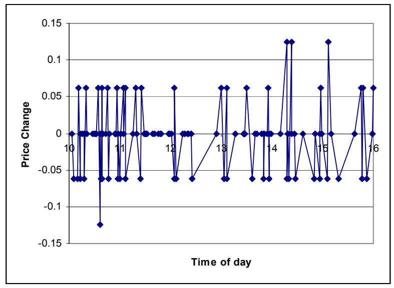
10

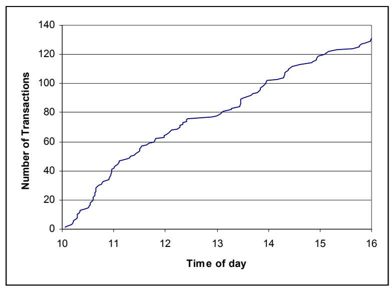
Econometric Framework

- All economic events can be described by when they occur and a list of characteristics.
- Sometimes the timing of the events is fixed, but with high frequency data it is not.
- For example,
  - Transaction events and prices
  - Limit order submission and strike
  - Quote revisions

Transaction Point Process







13

Notation

- Let *N(t)* denote the number of events that have occurred by time $t \in [0,T]$.
- Let t_i denote the i^{th} arrival time. These arrival times are referred to as a point process.
  - where $0=t_0<t_1<t_2<\ldots<t_{N(T)}=T$
- Let $x_i=t_i-t_{i-1}$ denote the i^{th} duration.
- Let y_i denote a vector of characteristics associated with the i^{th} arrival time.
- Jointly, the sequence of arrival times and marks are referred to as a *marked point process*.

14

Joint Density

- The joint density of the arrival times and the marks given the initial values for t_0 and y_0 is the object of interest.
- This can be recursively decomposed:

$$f\left(y_{N(T)}, y_{N(T)-1}, \ldots, y_1, t_{N(T)}, t_{N(T)-1}, \ldots, t_1 \mid y_0, t_0\right) = \prod_{i=1}^{N(T)} f\left(y_i, t_i \mid \tilde{y}_{i-1}, \tilde{t}_{i-1}\right)$$

where $\tilde{t}_i = \{t_i, t_{i-1}, \ldots, t_0\}$ and $\tilde{y}_i = \{y_i, y_{i-1}, \ldots, y_0\}$

15

How should such data be analyzed?

- Convert point process to fixed intervals and use traditional time series techniques.
  – Familiar ground – lots of available techniques.
- Model in tick time.
- Model as a point process.
  – When will the next event occur?
  – When will the next event of a specific type occur?
- Point process models for the marks. Joint modeling of prices and durations.

16

Fixed interval analysis.

- Most time series econometrics is based on fixed interval analysis.
- When the marks are of primary interest, there is a natural tendency to convert irregularly spaced data to fixed time intervals.
  - I.e. model prices over 5 minute intervals.

17

Numerous examples in the literature

- Andersen and Bollerslev (1998) estimate GARCH models for 5-minute FX returns constructed from the midquotes.
  - Sophisticated modeling of diurnal pattern.
  - Important to account for news announcements.

18

- Hasbrouck (2000), Zhang Russell and Tsay (2001) use 15 minute quote data to study quote dynamics.
- Hasbrouck(1995) (2001) use 1 second intervals.

19

Methods of converting to fixed intervals

- For some of the points in time defining the fixed intervals there will not be a corresponding event. What value should be used for the mark?

- Lets focus on the price as the mark.

20

Three main methods

Denote the log of the price at time t_i by p_i^*. Then let

$$p_t = \left(\lambda p_i^* + (1-\lambda)p_{i+1}^*\right) \text{where } t_i \le t < t_{i+1}$$

I. Use the **prevailing quote**: λ=1

II. **Interpolate**: $\lambda = \dfrac{t - t_i}{t_{i+1} - t_i}$

21

III. A third possibility is adopted by Hasbrouck (2002). Since time of a trade is usually recorded to the nearest second, then if the fixed interval is taken to be one second there is at most one observation per time period.

At each second the price is set either to this price or to the previous period price.

22

Prevailing price (λ=1): means and variances are not always retained.

Let the returns be denoted by:

$y_i{*} = p_i^* - p_{i-1}^*$ transaction time return (irregularly spaced)

$y_t = p_t - p_{t-1}$ interpolated value

- If there is never more than one trade per calendar interval then means and variances are preserved:

$$\sum_{i=1}^{N(T)} y_i{*} = \sum_{t=1}^{T} y_t, \text{ and } \sum_{i=1}^{N(T)} y_i{*}^2 = \sum_{t=1}^{T} y^2{}_t$$

- If the intervals are larger so that more than one trade can occur then means are preserved, but not variances.

$$\sum_{i=1}^{N(T)} y_i^* = \sum_{t=1}^{T} y_t, \text{ and } \sum_{i=1}^{N(T)} \left(\sum_{\substack{multiple\\trades}} y_i^* \right)^2 \neq \sum_{t=1}^{T} y_t^2$$

- If the high frequency returns are a martingale then the expectation of the cross products are zero and the expected value of the variances are the same.

- When prices are interpolated these relations no longer hold. The sum of the squared interpolated series will be:

$$\sum_{i=1}^{N(T)} \left[\tilde{y}_i\right]^2 = \sum_{i=1}^{N(T)} \left[\lambda_i p_i^* + (1-\lambda_i) p_{i-1}^* - \lambda_j p_j^* - (1-\lambda_j) p_{j-1}^*\right]^2$$

$$= \sum_{i=1}^{N(T)} \left[\lambda_i \left(p_i^* - p_{i-1}^*\right) + \left(p_i^* - p_j^*\right) + \left(1-\lambda_j\right)\left(p_j^* - p_{j-1}^*\right)\right]^2$$

where *i* and *j* are the events just after the two endpoints of the fixed interval

- Mean will be approximately right.
- If the returns form a martingale difference sequence then the expected variance and its probability limit will be less than the variance of the process.
- Autocorrelation is induced into the fixed interval returns.

Bivariate fixed interval analysis

- Consider the case that the fixed interval is set to the shortest interval over which the data can evolve (often this is one second).
  – Unlike aggregating to longer fixed intervals, no information is lost
  – If the prevailing price is used, lots of redundant values for the price are created. The implication is that many zero returns are created.
  – Does the “creation” of zeros induce bias?

Simple discrete time example

- Let z be the log price of an asset that is continuously observed at t=1,2,…,T. We assume the returns Δz_t are iid.
- Let y be the log price of a second asset that is observed at $N(T)$ random arrival times $t_1, t_2, \ldots, t_{N(T)}$ for t=1,2,…,T
- Let d_t denote an indicator for whether the price of y is observed at time t taking the value 1 with probability p.
- Then define the price at time t for asset y as

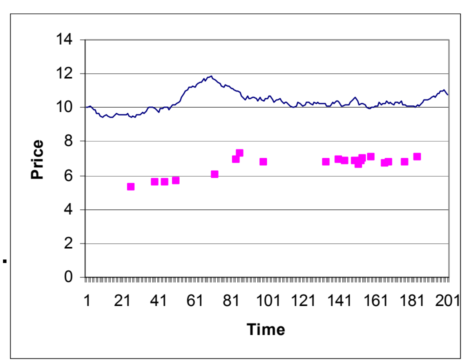
$$y_t = \begin{cases} y_{t-1} & \text{if N(t) = N(t-1)} \\ y_t & \text{if N(t) > N(t-1)} \end{cases}$$

- The return series for y will be zero if no price is observed at time t and will be non-zero when a price is observed at time t.

27

Goal: Regress Δy on Δz

- We consider two choices:
  - Regression using the returns associated with the N(T) random time intervals.
  - Regression using the T fixed interval returns.



28

Results:

- Suppose the true relationship between the two series is

$$\Delta y_{t_i} = \beta \Delta z_{t_i} + \varepsilon_i$$

or, equivalently $\Delta y_{t_i} = \beta \sum_{t_{i-1}}^{t_i} \Delta z_t + \varepsilon_i$

- If we use the T *fixed interval returns* and regress

$$\Delta y_t \text{ on } \Delta z_t$$

- Denote the estimate using the T fixed interval returns by $\hat{\beta}^T$

29

Conditional on whether the price for y is observed at time *t* yields the OLS estimator:

- then $\hat{\beta}^T = \left(\Delta z' \Delta z\right)^{-1} \sum_{t=1}^{T} \Delta z_t \left(\Delta z_t \beta + \varepsilon\right) d_t$

- Taking expectations of the conditional moments of the estimator yields:

$$p\lim\left(\hat{\beta}^T\right) = p\beta$$

- The resulting estimate $\hat{\beta}^T$ is downward biased. The bias is driven by the probability of observing the price of asset y.

30

Now suppose instead that we regress y_t on $z_t, z_{t-1}, \ldots, z_{t-K}$.

$$\hat{\beta}^T = \left(\Delta z' \Delta z\right)^{-1} \begin{bmatrix} \sum_{t=1}^{T} \Delta z_t \left(\Delta z_t \beta + \varepsilon\right) d_t \\ \sum_{t=1}^{T} \Delta z_{t-1} \left(\Delta z_{t-1} \beta + \varepsilon\right) d_t \left(1 - d_{t-1}\right) \\ \sum_{t=1}^{T} \Delta z_{t-2} \left(\Delta z_{t-2} \beta + \varepsilon\right) d_t \left(1 - d_{t-1}\right)\left(1 - d_{t-2}\right) \\ \vdots \\ \sum_{t=1}^{T} \Delta z_{t-k} \left(\Delta z_{t-k} \beta + \varepsilon\right) d_t \left(1 - d_{t-1}\right)\left(1 - d_{t-2}\right) \cdots \left(1 - d_{t-k+1}\right) \end{bmatrix}$$

31

Again, taking expectations over the conditional moment

- Then $p \lim\left(\hat{\beta}_k^T\right) = p\left(1-p\right)^{k-1} \beta$
- The estimates $\hat{\beta}_k^T$ decay exponentially in *p*.
- Additionally, for the infinite lag model $p \lim\left(\sum_{\lim K \to \infty} \hat{\beta}_k^T\right) = \beta$
- More generally, if p is not constant, the probability limit will be determined by

$$E_d\left(\left(\Delta z' \Delta z\right)^{-1} \sum_{t=1}^{T} \Delta z_{t-k} \Delta z_{t-k} d_t \prod_{j=1}^{k-1} \left(1 - d_{t-j}\right)\right) \beta$$

32

Interpretation is suspect however

- When we perform regression with the T equally spaced observations the result is slowly decaying parameter estimates.
- Interpreting this as long range dependence, or predictability is wrong. This long range dependence is purely an artifact of the estimation technique.
- Results are not useful for dynamic hedging or pricing of asset y.

33

Now suppose that z is measured randomly

- Probability of observing y in any period is p
- Probability of observing z in any period is q

$$E\left(\Delta y_t \Delta z_{t-k}\right) = pq\beta\sigma^2 E\left(overlap \left| \begin{matrix} \text{y observed at t} \\ \text{and z observed at t-k} \end{matrix} \right. \right)$$

$$E\left(\Delta y_t \Delta z_{t-k}\right) = pq\beta\sigma^2 \left(1-p\right)^k \frac{pq}{p+q-pq}$$

34

Regression Coefficients

- $$p\lim \beta_k = \beta\left(1-p\right)^k \; pq\,\frac{pq}{p+q-pq}$$

- and the sum of the coefficients in an infinite lag model is

$$\underset{K,T\to\infty}{p\lim} \sum_{k=1}^{K} \beta_k = \beta q \frac{pq}{p+q-pq}$$

35

IMPLICATIONS

- If q=1, then the sum of the coefficients will be the true response
- However, the apparent lag shape does not indicate market inefficiency or causality
- If p=1 but q<1, then the sum of the coefficients is q^2 so even the sum is understated
- If p<1 and q<1 then it is worse still.

36

When will the next event occur?

- Examples of Point Processes

- A point process can be described either in terms of the sequence of arrival times t_i or the sequence of durations x_i.
- Engle and Russell (1998) propose the Autoregressive Conditional Duration (ACD) to model the distribution of waiting times x_i conditional on the history of arrival times.
- Many point processes have been used in other fields of statistics

37

ACD

- The ACD model assumes that any dependence in the arrival rate can be summarized by:

$$\psi_i = E\left(x_i \mid \tilde{x}_{i-1}, \tilde{y}_{i-1}\right)$$

where

$$x_i = \psi_i \varepsilon_i$$

is iid with E(ε_i)=1 and $\tilde{z}_i \equiv \left\{z_i, .z_{i-1}, \dots .z_1\right\}$

38

Flexibility of the ACD

- The flexibility of the ACD model lies in the potential models for the mean ψ_i and the choice of the innovation distribution ε.
- Engle and Russell propose using the linear parameterization $\psi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}$
  – This is referred to as an ACD(p,q) since it contains p lags of *x* and q lags of ψ.
  – Often low order models such as the (2,2) model are capable of describing the temporal dependence in the arrival rates.

39

Interpretation of linear model

- Let $\eta_i = x_i - \psi_i$ which will be a margingale difference sequence by construction.
- Then the ACD(p,q) model can be written as an ARMA(p^*,q) model in the durations where p^*=max(p,q).

$$x_i = \omega + \sum_{j=1}^{p^*} \left(\alpha_j + \beta_j\right) x_{i-j} + \sum_{j=1}^{q} -\beta_j \eta_{i-j} + \eta_i$$

40

- Bauwens and Giot (2000), Russell and Engle (2004), Engle and Lunde (2003) consider the Nelson From ACD model.

$$\ln\left(\psi_i\right) = \omega + \sum_{j=1}^{p} \alpha_j \ln\left(\varepsilon_{i-j}\right) + \sum_{j=1}^{q} \beta_j \ln\left(\psi_{i-j}\right)$$

- Zhang Russell and Tsay (2003) consider a non-linear specification for the expected duration.

$$\psi_i = \begin{cases} \omega_1 + \alpha_1 x_{i-1} + \beta_1 \psi_{i-1} & \text{if } x_{i-1} \le a_1 \\ \omega_2 + \alpha_2 x_{i-1} + \beta_2 \psi_{i-1} & \text{if } a_1 < x_{i-1} \le a_2 \\ \omega_3 + \alpha_3 x_{i-1} + \beta_3 \psi_{i-1} & \text{if } a_2 < x_{i-1} \end{cases}$$

41

- Fernandez and Grammig (2003) consider a family of ACD models constructed from the Box-Cox Transformation.

42

Choice of the error term

- Suggestions for the distribution of ε_i include the exponential, weibul, (generalized) Gamma, and the Burr distributions.
- Implications of the choice of the distribution of ε_i are most easily seen by examining the conditional intensity function.

43

Conditional Intensity Function

Recall that *N(t)* denotes the counting function.
Then the conditional intensity function is given by:

$$\lambda\left(t \mid N(t), t_{i-1}, t_{i-1,\ldots,} t_0\right) = \lim_{\Delta t \to 0} \frac{\Pr\left(N\left(t+\Delta t\right) > N(t) \mid N(t), t_{i-1}, t_{i-1,\ldots,} t_0\right)}{\Delta t}$$

So the conditional intensity function characterizes the instantaneous probability of an event occurring given the history of the process.

That is, the probability that an event occurs over the next small time interval Δt is approximately given by

$$\lambda\left(t \mid N(t), t_{i-1}, t_{i-1,\ldots,} t_0\right) \Delta t$$

44

Baseline hazard function

45

- Let $p(\varepsilon;\phi)$ denote the density function of ε.
- Define the baseline hazard associated with ε as:

$$\lambda_0(\varepsilon) = \frac{p(\varepsilon;\phi)}{S(\varepsilon;\phi)}$$

where $S(\varepsilon;\phi) = 1 - \int_{\varepsilon}^{\infty} p(\varepsilon;\phi)$ is the survivor function.

Conditional Intensity for ACD

46

- From the baseline hazard $\lambda_0(\varepsilon) = \frac{p(\varepsilon;\phi)}{S(\varepsilon;\phi)}$ we obtain the conditional intensity function.
- Perform the change of variable $\varepsilon_{N(t)} = \frac{x_{N(t)}}{\psi_{N(t)}}$
- Then $\lambda\left(t \mid N(t), t_{i-1}, t_{i-1,\dots,} t_0\right) = \lambda_0\left(\varepsilon_{N(t)}\right) = \lambda_0\left(\frac{x_{N(t)}}{\psi_{N(t)}}\right)\frac{d\varepsilon}{dx} = \lambda_0\left(\frac{x_{N(t)}}{\psi_{N(t)}}\right)\frac{1}{\psi_{N(t)}}$
- Hence it is the shape of the distribution of ε that determines how the instantaneous probability of an event occurring evolves in the absence of a new event.

- For example, the exponential distribution implies the well known flat baseline hazard: $\dfrac{1}{\psi_{N(t)}}$

- The Weibull, on the other hand, allows for monotonic behavior of the hazard:

$$\lambda\left(t \middle| t_{N(t)},\ldots,t_0\right) = \left(\Gamma\left(1+\tfrac{1}{\gamma}\right)\psi_{N(t)+1}^{-1}\right)^{\gamma}\left(t - t_{N(t)}\right)^{\gamma-1}\gamma$$

- The Gamma and Generalized Gamma and Burr distributions allow for a rich class of hump shape hazards.

47

Diurnal Patterns

- Diurnal patterns in trading rates are well documented.
- We might consider formulating the conditional expectation as the product of a stochastic and a deterministic component

$$E_{N(t)-1}(x_{N(t)}) = \phi(t_{i-1};\theta_\phi)\psi_i(x_{i-1},..,x_1;\theta_\psi)$$

- Where $\phi(t_{i-1};\theta_\phi)$ is simply the expectation of the waiting time conditioned on the time of day that the duration starts.
- Two step and joint estimation are possible.

48

QMLE Results

- The EACD(1,1) model is clearly very similar to the GARCH(1,1) specification. The similarity is even closer than what you might think.
- Consider the likelihood for the EACD(1,1) model:

$$L = -\sum_{i=1}^{N(T)} \left(\log \psi_i + \frac{x_i}{\psi_i} \right)$$

- This is identical to the likelihood function for the GARCH(1,1) where $y_i = \sqrt{x_i}$
- Hence, QMLE results carry over from the GARCH literature (Lumsdaine (1996) or Lee and Hansen (1994)).
- If satisfies the conditions for y_i in their theorems then the EACD(1,1) model is a QMLE.

49

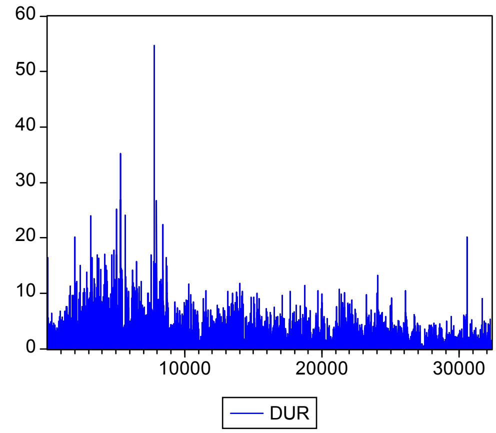
QMLE Implications

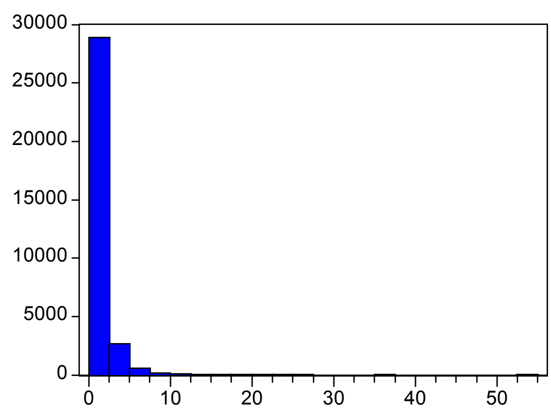
- We can use standard GARCH software to estimate parameters (ω and the α's and β's) of the conditional expected duration for the linear ACD(1,1) model.
- Given consistent estimates of these parameters one can non-parametrically estimate the baseline hazard model using:

$$\hat{\varepsilon}_i = \frac{x_i}{\hat{\psi}_i}$$

50

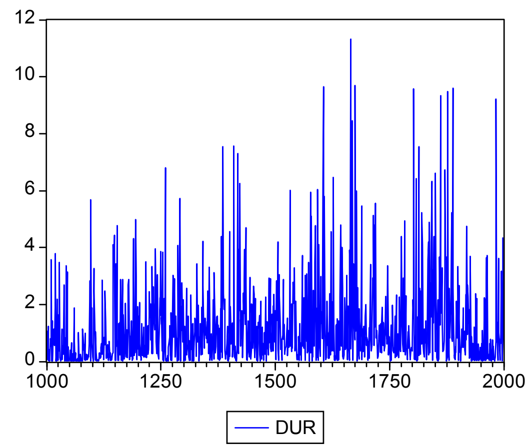
Airgas Stock Example



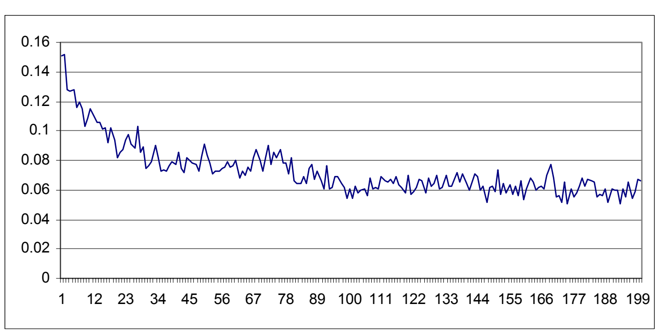


Series: DUR
Sample 1 32366
Observations 32366

Mean	0.996958
Median	0.444078
Maximum	54.66090
Minimum	0.005969
Std. Dev.	1.574711
Skewness	5.236681
Kurtosis	77.04077
Jarque-Bera	7540910.
Probability	0.000000



ARG duration autocorrelations



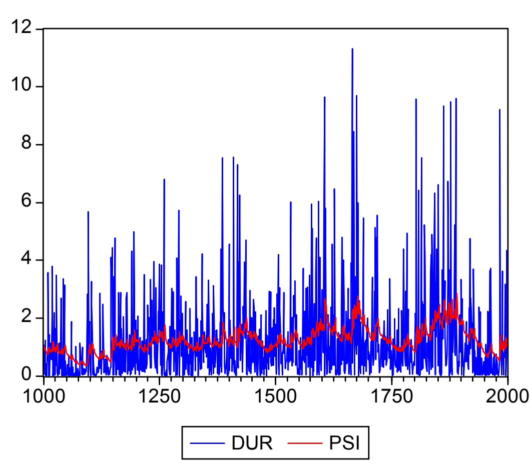
ACD Estimates for ARG

Example using GARCH estimation code

	Coefficient	Robust Std. Err.
ω	0.004244	0.000855
α_1	0.070261	0.007157
α_2	0.038710	0.012901
α_3	-0.055966	0.008640
β_1	0.835806	0.125428
β_2	0.107894	0.118311

55

Observed and Predicted Durations



56

Model Evaluation

- Any inhomogeneous Poisson process can be converted to a homogeneous Poisson process by a deterministic transformation of the time scale (see for example Snyder and Miller Springer).

$$u_i = \int_{s=t_{i-1}}^{t_i} \lambda\left(s \mid N(t), t_{i-1}, t_{i-1,\ldots,} t_0\right) ds$$

The "durations" measured on the u_i time scale should be a homogeneous Poisson with unit intensity.
Clearly this can be tested given an estimated model.

57

Example

- For the exponential model:

$$u_i = \int_{s=t_{i-1}}^{t_i} \lambda\left(s \mid N(t), t_{i-1}, t_{i-1,\ldots,} t_0\right) ds = \int_{s=t_{i-1}}^{t_i} \frac{1}{\psi_{N(t)}} ds = \frac{t_i - t_{i-1}}{\psi_{N(t)}} = \frac{x_i}{\psi_{N(t)}}$$

- For the Weibull model:

$$u_i = \int_{s=ti_{-1}}^{t_i} \lambda\left(t \middle| t_{N(t)}, \ldots, t_0\right) = \int_{s=t_{i-1}}^{t_i} \left(\Gamma\left(1+{}^{1}\!/_{\gamma}\right)\psi_{N(t)+1}^{-1}\right)^{\gamma} \gamma\left(s - t_{i-1}\right)^{\gamma-1} ds$$

$$= \left(\Gamma\left(1+{}^{1}\!/_{\gamma}\right)\psi_{N(t)+1}^{-1}\right)^{\gamma} \left(t_i - t_{i-1}\right)^{\gamma} = \left(\Gamma\left(1+{}^{1}\!/_{\gamma}\right)\frac{x_i}{\psi_{N(t)+1}}\right)^{\gamma}$$

58

Diagnostics

- Are the $\hat{u}_i$ uncorrelated?
  - Perform serial correlation test such as Ljung Box
- Are they approximately unit exponential?

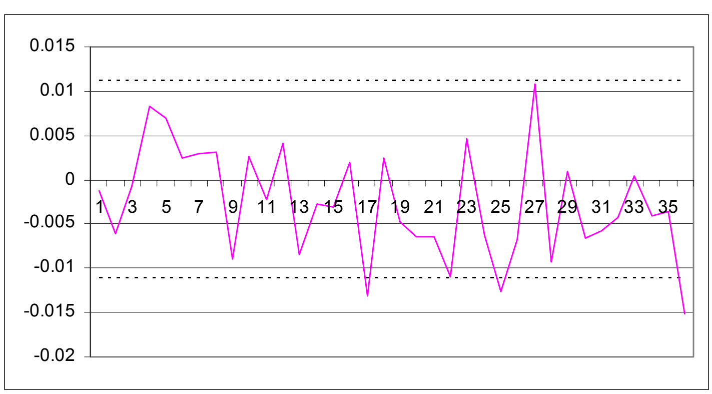
$$N(T)\frac{sd(u)-1}{\sqrt{8}}$$

should have a limiting standard Normal distribution.

- Alternatively, out of sample prediction could be examined or predictive distributions (Diebold Gunther and Tay (1998)

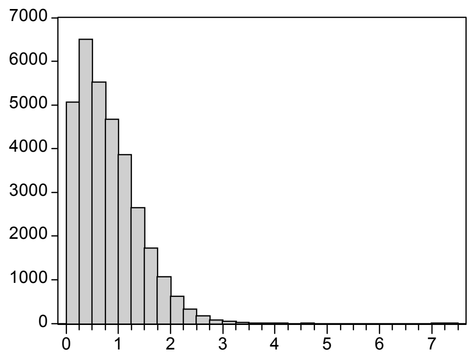
59

Residual Autocorrelations for ARG



60

Does it look like the Exponential distribution assumption is valid?



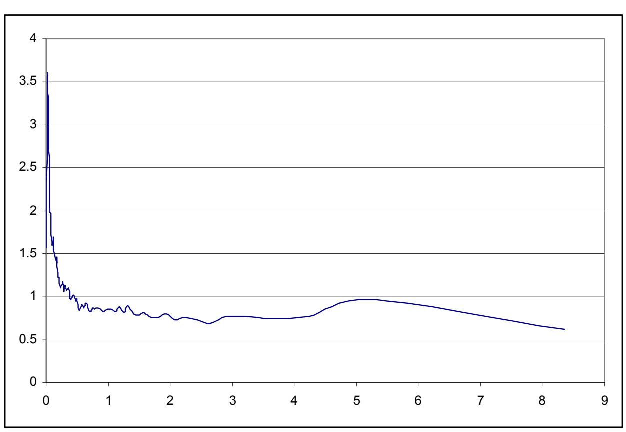
Series: EHAT
Sample 1 32366
Observations 32366

Mean	0.821029
Median	0.704446
Maximum	7.253621
Minimum	0.031156
Std. Dev.	0.570596
Skewness	1.115027
Kurtosis	5.046479
Jarque-Bera	12354.67
Probability	0.000000

$$\sqrt{32366}\,\frac{\sigma_u - 1}{\sqrt{8}} = \sqrt{32366}\,\frac{.57 - 1}{\sqrt{8}} = -27$$

61

Non-Parametric Estimate of Baseline Hazard



62

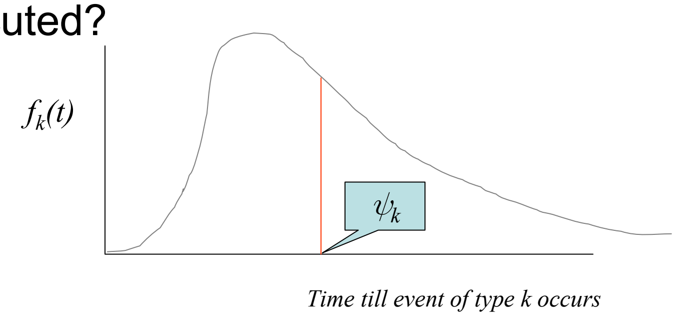
There may be several types of events

- Model only these events (thinning)
- Build a joint model to determine the arrival probabilities of different types of events (more later).

63

When will the next event of a particular type occur?

- Prices: How long will it take for the price to move more than an amount *c*?
- Execution: How long until the a transaction involving a limit order is executed?



64

Modeling the marks in *Tick Time*.

- Models the marks y_i as a time series where i indexes the i^{th} event arrival.
- Since there is no aggregation to fixed intervals no information is lost.
- This modeling approach has proven useful in the analysis of a single stock. The model operates in the time scale that new information occurs.

65

Examples include Hasbrouck 1991

- VAR's used to model the bivariate system of prices and volume for a given stock.
- Where m_i is the prevailing price (defined as the midpoint of the bid and ask) at the time of the i^{th} trade.
- w_i is the volume associated with the i^{th} transaction.

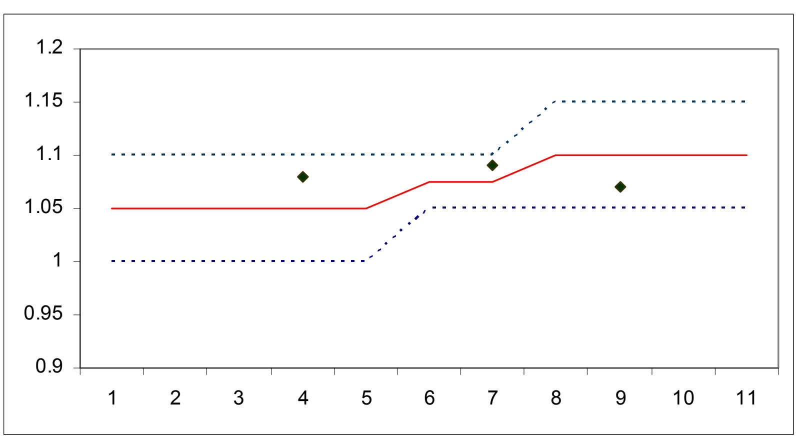
$$m_i = \sum_j a_j m_{i-j} + \sum_j b_j w_{i-j} + \varepsilon_{1i}$$
$$w_i = \sum_j c_j m_{i-j} + \sum_j d_j w_{i-j} + \varepsilon_{2i}$$

66

What is the expected price impact of a trade?

- Since order flow is correlated it makes more sense to ask "what is the expected price impact of the unexpected portion of a trade?".
- Hasbrouck argues that the market structure dictates a specific ordering in the impulse response functions. Namely, he argues, that market orders hit existing prices in the market.

67

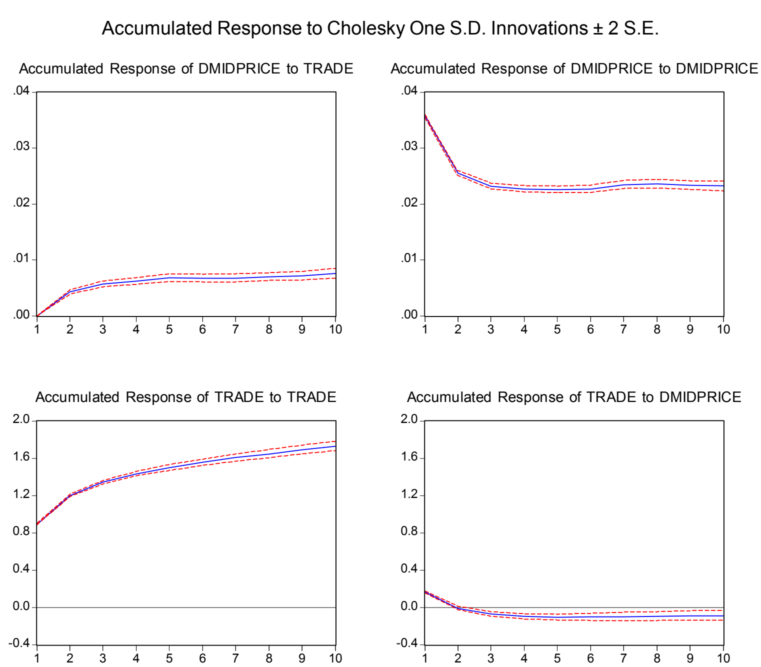


68

VAR for Airgas

- We estimate a VAR for midquote prices and a buy/sell indicator.
- Needs lots of lags (about 10).

69



- Tick time models are difficult to apply to multivariate data. For example, consider joint modeling of two stock price return series.
- Two different time scales, how should they be combined?
- Fixed interval models clearly have an advantage here in that once fixed interval data is obtained all the usual econometrics tools (VARs etc) can be applied.
- Finally, if the properties of the time series depend on the spacing of the data the tick time models may be mispecified.

71

Marked point process approach

- Without loss of generality we can always decompose the joint distribution into the product of a conditional and a marginal:

$$f\left(y_i, t_i \mid \tilde{y}_{i-1}, \tilde{t}_{i-1}; \theta\right) = \underbrace{g\left(y_i \mid \tilde{y}_{i-1}, \tilde{t}_i \mid \theta_1\right)}_{\text{Distribution of the Mark}} h\left(t_i \mid \tilde{y}_{i-1}, \tilde{t}_{i-1} \mid \theta_2\right)$$

- Here *g* denotes the conditional density of the mark given historical information as well as the contemporaneous duration

72

Interesting Marks Considered

- Transaction price returns.
  - Engle (2000), Ghysels and Jasiak (1998) GARCH models.
  - Russell and Engle (2004), Rydberg and Shephard (2000) and many other recent contributions model discrete transaction price moves.
- Order type (limit, market, marketable limit order, cancellation etc…).

73

Ultra-High Frequency Volatility Models

- Engle (2000) proposes a GARCH model for transactions data.
- The idea is that the volatility per unit time follows a GARCH process. The volatility per trade will likely depend on the time interval.

74

UHF GARCH set up

- Let r_i denote the return from transaction *i*-1 to transaction *i*.
- Denote the volatility per trade by: $h_i = Var\left(r_i \mid \tilde{x}_i, \tilde{r}_{i-1}\right)$
- Denote the volatility per unit time by:

$$\sigma_i^2 = Var\left(\frac{r_i}{\sqrt{x_i}} \mid \tilde{x}_i, \tilde{r}_{i-1}\right)$$

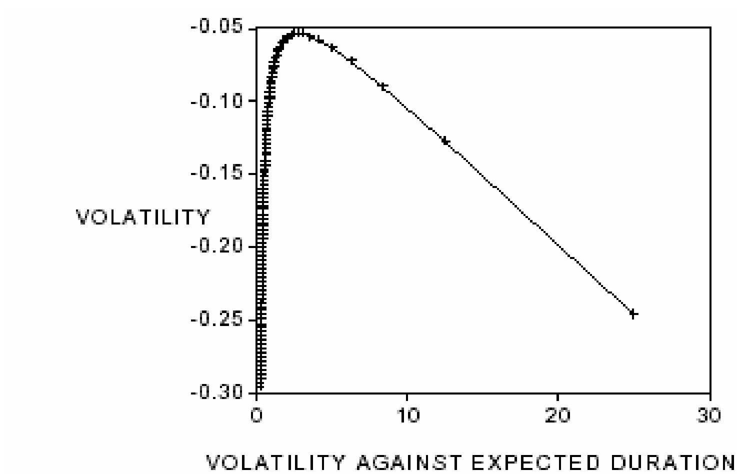
75

- The volatility per unit time is hypothesized to follow a GARCH process.
- After modeling the mean return of r_i let e_i denote the innovation.
- Then $\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma\left(\tilde{x}_i\right)$
- Using joint ACD model Engle proposes

$$\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma_1 x_{i-1}^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \psi_i^{-1} + \gamma_4 \xi_i$$

Long run volatility measure
exponential smoothing

76



77

Models for discrete price changes

- Russell and Engle (2004) propose the ACM model for discrete price changes.
- When used jointly with the ACD model for the durations it is referred to as the ACM-ACD model

78

SPECIFYING THE PROBABILITY STRUCTURE

LET $\tilde{x}_t$ and $\tilde{\pi}_t$ be the kx1 vectors indicating the state observed and the conditional probability of all k states respectively.

That is, $\tilde{x}_t$ takes the j^{th} column of the kxk identity matrix if the j^{th} state occurred.

A first order markov chain
(1) $\tilde{\pi}_t = P\tilde{x}_{t-1}$
links these with a transition probability matrix *P* with the properties that

a) all elements are non-negative
b) all columns sum to unity

79

In a more general setting *P* will be the conditional transition matrix and will vary with information available at time t-1. In this context this will include longer lags on x , π , and the time since the last transaction as well as other parameters of the timing of trades, and economic variables such as spreads, volume and other measures of market liquidity.

The restrictions on *P* are directly satisfied by simple estimators in the case of a constant transition matrix but are difficult to impose in simple linear extensions.

80

Here we propose an inverse logistic transformation which imposes such conditions directly for any set of covariates.

$$\log(\tilde{\pi}_{im} / \tilde{\pi}_{ik}) = \log\left(\sum_{j=1}^{k} P_{mj}\tilde{x}_{(i-1)j}\right) - \log\left(\sum_{j=1}^{k} P_{kj}\tilde{x}_{(i-1)j}\right)$$
$$= \sum_{j=1}^{k} \log\left(P_{mj} / P_{kj}\right)\tilde{x}_{(i-1)j}$$
$$= \sum_{j=1}^{k-1} P^{*}_{mj} x_{(i-1)j} + c_m$$

81

Rewriting the *k*-1 log functions as h() this can be written in simple form as:

$$h(\pi_i) = P^{*} x_i + c$$

where P^* is an unrestricted (k-1)x(k-1) matrix *c* is an unrestricted (k-1)x1 vector and x is a the (k-1)x1 state vector.

82

From estimates of P^* and the vector c, we find that

$$P_{mn} = \frac{\exp[P^*_{mn} + c_m]}{1 + \sum_{j=1}^{k-1} \exp[P^*_{jn} + c_j]}$$

so that all probabilities are positive including the probabilities of state k which are obtained from condition b).

83

Now by generalizing to allow for more dynamics, we are generalizing the transition matrix to allow the conditional transition probabilities to vary. For a first order model with predetermined or weakly exogenous variables z that will generally contain a constant,

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

84

An expression for the probability of observing a state can similarly be expressed in terms of the past history of the process:

$$\frac{\pi_i}{1-\iota'\pi_i} = \exp\left[P^* x_{i-1} + c\right]$$

$$\pi_i = \frac{\exp[P^* x_{i-1} + c]}{1+\iota'\exp\left[P^* x_{i-1} + c\right]}$$

where exp[P*] is interpreted as a matrix with elements $\exp[P^*_{mn}]$, and ι is a vector of ones.

85

More generally, we define the Autoregressive Conditional Multinomial (ACM) model as:

$$\boxed{h(\pi_i) = \sum_{j=1}^{p} A_j\left(x_{i-j} - \pi_{i-j}\right) + \sum_{j=1}^{q} B_j h\left(\pi_{i-j}\right) + \chi Z_i}$$

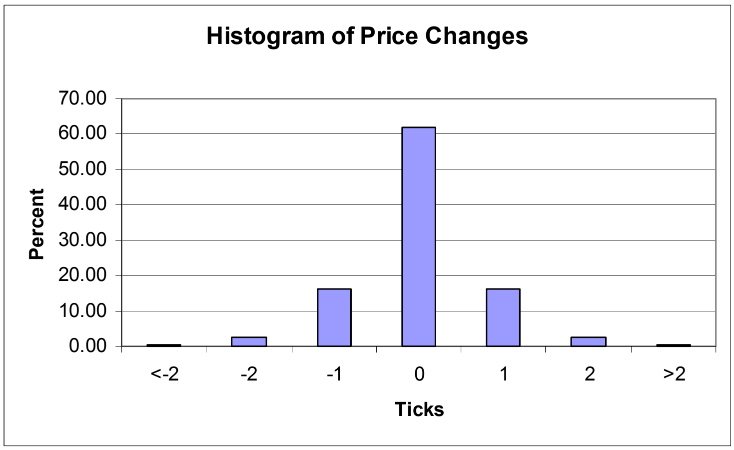
Where $h(\cdot):(K-1)\rightarrow(K-1)$ is the inverse logistic function.

Z_i might contain t_i, a constant term, a deterministic function of time, or perhaps other weakly exogenous variables.

We call this an ACM(p,q) model.

86

Lets consider the ARG transaction price changes.



87

We therefore consider a 5 state model defined as

$$x_i = \begin{cases} [1,0,0,0]' & \text{if } \Delta p_i \leq \text{-2 ticks} \\ [0,1,0,0]' & \text{if } \Delta p_i = -1 \text{ tick} \\ [0,0,0,0]' & \text{if } \Delta p_i = 0 \\ [0,0,1,0]' & \text{if } \Delta p_i = +1 \text{ tick} \\ [0,0,0,1]' & \text{if } \Delta p_i \geq +2 \text{ ticks} \end{cases}$$

It is interesting to consider the sample cross correlogram of the state vector x_i.

88

Sample cross correlations of x

$$
\begin{array}{ll}
 & s = 1 \\
\begin{array}{r} \text{up 2} \\ \text{up 1} \\ \text{down 1} \\ \text{down 2} \end{array} &
\begin{bmatrix} - & - & + & + \\ - & - & + & + \\ \cdot & + & - & - \\ + & + & \cdot & - \end{bmatrix} \\
 & \begin{array}{cccc} \text{up 2} & \text{up 1} & \text{down 1} & \text{down 2} \end{array}
\end{array}
\quad
\overset{2}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & + & + \\ \cdot & + & \cdot & \cdot \\ + & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{3}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & + & \cdot & \cdot \\ \cdot & \cdot & + & \cdot \\ + & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{4}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & + & \cdot & \cdot \\ \cdot & \cdot & + & \cdot \\ + & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{5}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & + & \cdot \\ + & \cdot & \cdot & + \end{bmatrix}}
$$

$$
\overset{6}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & + & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{7}{\begin{bmatrix} + & - & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{8}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{9}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & + & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ + & + & \cdot & + \end{bmatrix}}
\quad
\overset{10}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & + & \cdot \\ \cdot & \cdot & \cdot & + \end{bmatrix}}
$$

$$
\overset{11}{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ + & \cdot & + & + \end{bmatrix}}
\quad
\overset{12}{\begin{bmatrix} \cdot & \cdot & \cdot & + \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & \cdot \end{bmatrix}}
\quad
\overset{13}{\begin{bmatrix} + & \cdot & \cdot & + \\ \cdot & \cdot & - & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & + \end{bmatrix}}
\quad
\overset{14}{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}}
\quad
\overset{15}{\begin{bmatrix} \cdot & + & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & + & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}}
$$

Are there deterministic patterns in the price movements?

- Deterministic patterns in fixed interval volatility.
- Deterministic patterns in durations.
- Stochastic volatility has been found to be explained by stochastic transaction rates.
- Related question is whether deterministic patterns in transaction rates are driven by large per trade price changes or simply faster trading.

Deterministic Regression Results

	Const.	d_1	d_2	d_3	d_4	d_5	d_6	*F stat p-value*
Durations	239.30 (11.83)	-16.84 (17.02)	65.20 (14.24)	43.08 (15.03)	-41.88 (15.10)	-63.26 (14.39)	-118.25 (17.69)	0.00%
Down 2	.030 (.0053)	-0.0034 (.0076)	0.0083 (.0064)	-0.0000 (.0067)	-.0153 (.0067)	0.00625 (.0064)	0.00818 (.00795)	6.432%
Down 1	0.165 (.0117)	-.0077 (.0168)	.0060 (.0140)	-.0046 (.0148)	.00659 (.0149)	.00773 (.0142)	-.0292 (.0175)	61.3%
Up 1	.156 (.0117)	.0147 (.0169)	-.0106 (.0141)	.0196 (.0149)	-.029 (.0149)	.0030 (.0142)	.0227 (.0175)	25.6%
Up 2	.0395 (.0052)	-.0172 (.0075)	.0106 (.0063)	-.0042 (.0066)	-.0059 (.0067)	.0053 (.0064)	.0013 (.0078)	23.3%

91

- We propose the following model for the ACM

$$h(\pi_i) = c + \sum_{j=1}^{p} A_j \left(x_{i-j} - \pi_{i-j}\right) + \sum_{j=1}^{q} B_j h(\pi_{i-j}) + \sum_{j=1}^{r} \chi_j \ln(\tau_{i-j+1})$$

- The ACD is assumed to follow the Nelson form ACD with exponential error
- past price changes potentially influencing future durations:

$$\ln(\psi_i) = \omega + \sum_{j=1}^{u} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{v} \beta_j \ln(\psi_{i-j}) + \sum_{j=1}^{w} \left(\rho_j y_{i-j} + \zeta_j y_{i-j}^2\right)$$

92

Likelihood

- The log likelihood for the ACM part looks like

$$L = \sum_{i=1}^{N} \sum_{j=1}^{K} \left(\tilde{x}_{ij} \log\left(\tilde{\pi}_{ij}\right) \right) = \sum_{i=1}^{N} \tilde{x}_i' \log\left(\tilde{\pi}_i\right)$$

- Of course, the joint likelihood for the ACM-ACD model is obtained by summing the ACM and the ACD log likelihoods.
- The recursive structure of the model permits closed form evaluation of the likelihood function subject to initial conditions.

Several models are estimated in the paper. A simple to general model selection procedure suggests an ACM(3,3,3)-ACD(2,2). There are a lot of parameters estimated so I won't show them here.

$$v^*_i = x_i - \hat{\pi}_i$$

$$v_i = U_i^{-1} v^*_i$$

where $U_i V_i U_i' = I$

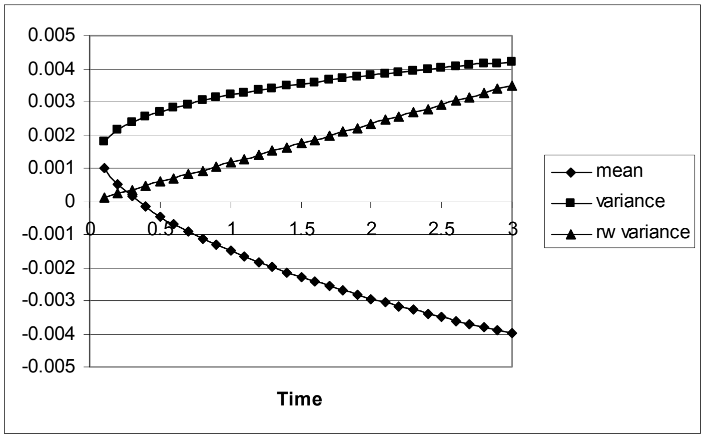
$$E\left(v_i \middle| I_{i-1}\right) = 0$$

$$E\left(v_i v_i' \middle| I_{i-1}\right) = \mathrm{I}$$

LB: x_i=23,324
LB: v_i=423

$$\begin{array}{ccccc}
\text{s = } 1 & 2 & 3 & 4 & 5 \\
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} + & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & + \end{bmatrix} \\
6 & 7 & 8 & 9 & 10 \\
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & + & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & + & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & - & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} + & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & - & \cdot \\ \cdot & \cdot & + & + \\ \cdot & \cdot & \cdot & + \\ \cdot & + & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \\
11 & 12 & 13 & 14 & 15 \\
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} - & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & - & - & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} &
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & + & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}
\end{array}$$

Implied relationship between trade intervals and price changes.



95

Concluding Remarks

- Three general approaches to analyzing high frequency financial data.
  - Point process
  - Tick time
  - Fixed interval
- Choice of approach is driven by the goal/question
- Relationship between approaches is underdeveloped.
  - Biases may be present
  - Interpretation may be suspect.

96

Measuring Transactions Cost

- Market microstructure seeks to understand the workings of financial markets.
- One of the most fundamental features of a market is the quality of execution of orders.
- An ideal market is one in which market participants can transact as costlessly as possible.

How should we measure transaction costs?

- Ideally, we would like to examine transaction costs that allow for a variety of trading strategies.
- There is a tradeoff between the cost of immediacy and the risk of patiently trading over a longer period of time.
  - Typically, the larger the quantity traded at once, the worse the price obtained.
  - Breaking the trade up into small chunks decreases the expected cost of the trade, but exposes the trader to risk of movements in the underlying.
- Any measure of transaction cost, however, necessarily requires measuring the cost of any individual transaction

What is the cost of a single trade?

- Bid-ask spreads are one measure of the cost of a single trade. However, in many markets there is room for price improvement so that trades often occur strictly inside the bid ask spread.
- An alternative measure that accounts for price improvement is the *effective cost of trade*.

Notation

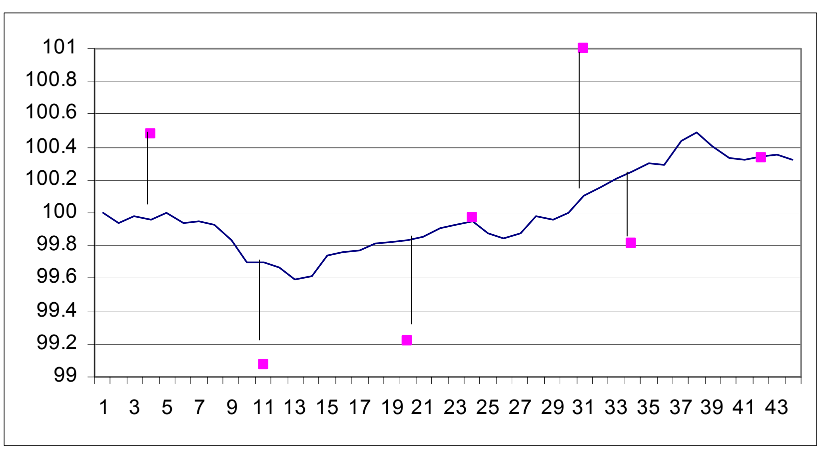
- Let m_t denote the “efficient price”. This is the price that would prevail in equilibrium in absence of any market frictions (ie if at any point in time there were a single price at which transactions occur).
- Let p_i denote the i^{th} transaction price.
- Let Q_i denote an indicator for whether the i^{th} trades is a market buy or sell order taking the values 1 and -1 respectively.

- Then the effective spread is defined as:

$$E\frac{\left[Q_i\left(p_i - m_{t_i}\right)\right]}{m_{t_i}}$$

- Problems:
  - only in rare data sets to we observe Q_i.
  - We don't observe the efficient price.

- Solution
  - Use an algorithm to assign trades as buyer or seller initiated.
  - Use midpoint of bid ask to proxy for efficient price.



The vertical distances represent the cost to the trader.

Roll's measure

- Let the returns be given by: $p_{t_i} = m_{t_i}\eta_{t_i}$

Pr(Q_i=1)=Pr(Q_i=-1)=.5 and Q_i is iid.

Further assume that $\ln\left(\eta_{t_i}\right) = \frac{s}{2}Q_{t_i}$

Then the return is given by:

$$\underbrace{\ln\left(p_{t_i}\right) - \ln\left(p_{t_{i-1}}\right)}_{r_i} = \underbrace{\ln\left(m_{t_i}\right) - \ln\left(m_{t_{i-1}}\right)}_{r_i^e} + \underbrace{\left(\frac{s}{2}Q_{t_i}\right) - \left(\frac{s}{2}Q_{t_{i-1}}\right)}_{\varepsilon_i}$$

$$r_i = r_i^e + \varepsilon_i$$

If we additionally assume that the *efficient price follows a martingale difference sequence* and is *uncorrelated with the noise return* we get:

$$\operatorname{cov}\left(r_i, r_{i-1}\right) = E\left[\left(r_i^e + \varepsilon_i\right)\left(r_{i-1}^e + \varepsilon_{i-1}\right)\right] = E\left(\varepsilon_i \varepsilon_{i-1}\right)$$

$$= E\left[\left(\left(\frac{s}{2}Q_i\right) - \left(\frac{s}{2}Q_{i-1}\right)\right)\left(\left(\frac{s}{2}Q_{i-1}\right) - \left(\frac{s}{2}Q_{i-2}\right)\right)\right]$$

$$= -E\left(\frac{s^2}{4}Q_{i-1}^2\right) = -\frac{s^2}{4}$$

So $s = 2\sqrt{-\operatorname{cov}\left(r_i, r_{i-1}\right)}$

Summary of Roll's measure

- Doesn't require a reference price like the efficient price – only requires transaction prices.
- Can be estimated based on daily data.

However

- Sometimes inconsistent with the data: positive covariances are often obtained.
- Strong assumptions regarding dependence structure of efficient price and noise process.

- Bandi and Russell (2004) propose an estimator that, like Roll's model only requires transactions prices.
- Unlike Roll's model there can be arbitrary dependence in the cost dynamics and unrestricted dependence between the efficient price and the cost of trade.

Russell Tsay and Zhang (2002) Econometric Modeling Goals

- Model the discrete bid and ask prices.
- Time varying volatility
- Time varying liquidity
- Diurnal patterns

- Address economic questions regarding price discovery and liquidity.

Time series models

- Bollerslev and Melvin (1994) propose modeling FX data via a GARCH process and feeding the GARCH volatilities into an ordered probit model for the bid ask spread.
- Engle and Patton (2000) propose an error correction model for the bid and ask prices where the spread is the correction variable.

Decomposition models for discrete bid and ask prices

Let m_t denote the log of the "true" efficient price.

Let a_t and b_t denote the observed ask and bid price.

Let $\alpha_t > 0$ and $\beta_t > 0$ denote the cost of exposure on the ask and bid side respectively.

$$m_t = m_{t-1} + v_t \qquad v_t \sim N(0, \sigma_t^2)$$

$$a_t = round^a\left(m_t + \alpha_t\right)$$

$$b_t = round^b\left(m_t - \beta_t\right)$$

Interpretations of the cost

- I will refer to the "cost of bid exposure" and the "cost of ask exposure".
- On the NYSE the specialist chooses bid and ask prices at which a maximum quantity can be traded.
- Hence the cost of exposure is the amount the specialist is compensated for fixed cost and risk.
  - Large cost low liquidity and vice versa.
  - Motivation in other studies to consider the spread as a measure of liquidity.

Special Cases

- If $\alpha_t = \beta_t = c$, $\sigma_t^2 = \sigma^2$, and *round*^{a}=*round*^{b}=*round* to the nearest tick we get the model of Harris (1990)

- If there are no costs (c=0) and *round*^{a}=*round*^{b}=*round* to the nearest tick we get the model for transaction prices proposed by Roll (1984)

- These simple models focused on the effects of discrete price observations on volatility estimates for the efficient price.

Time varying cost and volatility

- More generally, we would like to let the cost functions α_t and β_t be time varying denoting variability in the cost functions faced by traders.

- A realistic model should also allow for time varying volatility (or GARCH effects).

Our Model

Rounding

$$b_t = Floor(m_t - \beta_t)$$
$$a_t = Ceiling(m_t + \alpha_t)$$

Cost

$$\ln(\alpha_t) = \mu_t + \theta_t + \theta_t^{\alpha} + \phi(\ln(\alpha_{t-1}) - \mu_{t-1}) + \nu_t^{\alpha}$$
$$\ln(\beta_t) = \mu_t + \theta_t + \theta_t^{\beta} + \phi(\ln(\beta_{t-1}) - \mu_{t-1}) + \nu_t^{\beta}$$

Efficient Price

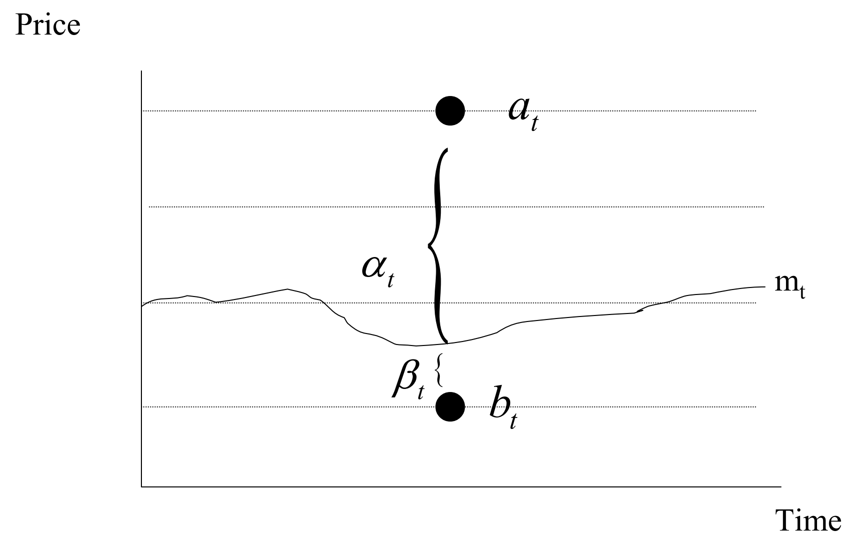
$$m_t = m_{t-1} + \delta_t + \varepsilon_t \qquad \varepsilon_t = \sigma_t z_t$$
$$\ln(\sigma_t^2) = \eta_t + \varsigma_t + \varphi(\ln(\sigma_{t\ t}^2) - \eta_{t-1}) + \gamma z_t$$

$$z_t \sim iid\ GED$$

μ_t and η_t are deterministic functions of time of day.
θ_t are variables with a common impact on both cost functions
θ_t^{α} and θ_t^{β} are variables that affect the ask and bid sides only respectively
δ_t is the impact of variables on the efficient price
$\nu_t^{\alpha}, \nu_{\mathrm{t}}^{\beta}$, and z_t are iid variables and mutually independent

Advantages of Decomposition

- The decomposition models have the advantage of allowing for, and potentially providing quantitative measures of separate cost functions for the bid side and the ask side (the spread measures the sum of the two cost functions plus rounding noise)
- The permanent impact of trade characteristics on the efficient price m_t can be assessed. (price discovery process)
- We can estimate a volatility model for the efficient price (not contaminated by discrete measurement).



Related model

- Hasbrouck (1999a, 1999b) proposes and estimates a similar model. His goal is to show that the cost functions are asymmetric and time varying both stochastically and deterministically.
- Here we are interested in how characteristics of the market impact the dynamics of the cost functions and the efficient price.
- In doing so we will learn about market liquidity and price discovery. We pay particular attention to the effects of order flow.

Estimation

θ_1 Parameters of the cost model

θ_2 Parameters of the Nelson EGARCH model

s State vector of unobserved components including m_t, α_t, and β_t.

The likelihood involves muti-dimensional integration which must be solved numerically. We follow Manrique and Sheppard (1997) and Hasbrouck (1999b) and use MCMC methods with uninformative priors.

Estimation details are in the appendix of the paper.

The data

- We extract trades from the TORQ data set spanning the three
- Months Nov 90 to Jan 91.
- 84.6% of the bid and ask changes are of 1 tick or are unchanged.
- For tractability we follow Hasbrouck and use quotes observed just prior to the end of fixed 15 minute intervals when the model is estimated.

Market Data

- **LogSpread** = log of mean bid-ask spreads over 15 minute interval
- **LogTPriceVar**= log of variance of trade by trade prices over interval
- **LogVolume+=** log of cumulative positively signed volume over interval
- **LogVolume-**= log of cumulative negatively signed volume over interval
- **LogAskDepth**= log of mean ask depth over interval
- **LogBuyDepth**= log of mean buy depth over interval

- The signed volume is obtained using the Lee and Ready rule.

Processing of Market Data

- All market variables contain time of day effects.
- All of the market variables are highly persistent (autocorrelated).
- To aid interpretation of the model it is useful to decompose the market variables into their predictable component and the “surprise” or innovation.
- The quote updates may reflect only the surprise element of the market variable.
- In this spirit after subtracting off its deterministic component we decompose each market variable into its predictable and surprise components using ARMA(p,q) models.

- We are left with expected and unexpected components of each of the market variables.
- We define two more volume variables:
  - Level: total unanticipated=surprise buyer initiated volume + surprise seller initiated volume.
  - Pressure: Unanticipated order imbalance=surprise buyer initiated volume – surprise seller initiated volume.
  - |Pressure|

Specification

•Deterministic components:

$$\mu_t = k_1 + k_2^{open} \exp(-k_3^{open} \tau_t^{open}) + k_2^{close} \exp(-k_3^{close} \tau_t^{close})$$

$$\eta_t = \eta^{night} \left(I_{\tau_t^{open}} = 0 \right) + l_1 + l_2^{open} \exp(-l_3^{open} \tau_t^{open}) + l_2^{close} \exp(-l_3^{close} \tau_t^{close})$$

The GED tail parameter can be different overnight as well.

- **Cost equations** contain expected and unexpected components as well as the level and pressure variables.
- The **volatility equation** contains expected and unexpected components.
- The **efficient price** dynamics contain unexpected components only. It is therefore unforcastable and remains a random walk.

Expectations

- The cost equations denote compensation to the specialist for exposure to trade.
  - Volume effects mixed.
  - Volatility increased risk => increased cost of exposure.
- Likely order imbalance will be most informative for the direction of movements of the efficient price.
- Volume and spreads should be informative about the volatility of the efficient price.

Model Building

- Starting with both the predicted and unexpected components entering unrestricted into the two cost models and the EGARCH model we use a general to simple model selection approach.
- The predicted components drop out with the exception of depth with has a significant coefficient on the predictable component, but insignificant unexpected component.

Symmetry in the Cost Functions

- Symmetry in the bid and ask dynamics are tested.
- Transaction price volatility enters symmetrically in cost.
- Signed volume effects are asymmetric in cost models.

Some Model Diagnostics

- Somewhat extensive model diagnostics presented in the paper suggest no additional lags are needed in the cost model but 2 lags are needed in the volatility model.

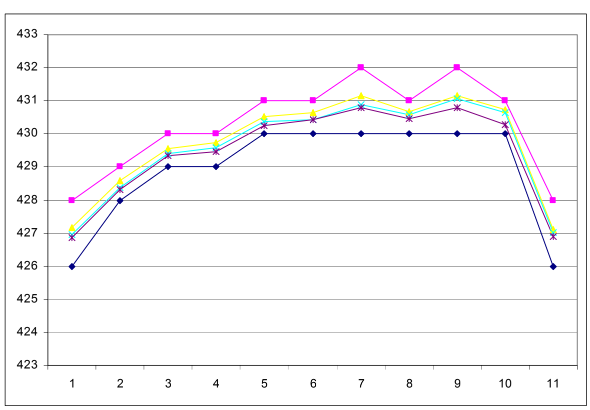
Table V: **Parameter Estimates of the Extended Hasbrouck Model for GE**

$$
\left\{
\begin{array}{rcl}
a_t & = & \lceil m_t + \alpha_t \rceil \\
b_t & = & \lfloor m_t - \beta_t \rfloor \\
\\
\ln(\alpha_t) & = & \mu_t + \theta_t + \theta_t^\alpha + \phi[\ln(\alpha_{t-1}) - \mu_{t-1}] + \sigma_v v_t^\alpha \\
\ln(\beta_t) & = & \mu_t + \theta_t + \theta_t^\beta + \phi[\ln(\beta_{t-1}) - \mu_{t-1}] + \sigma_v v_t^\beta \\
\mu_t & = & k_1 + k_2^{open} \exp\left(-k_3^{open} \tau_t^{open}\right) + k_2^{close} \exp\left(-k_3^{close} \tau_t^{close}\right) \\
\theta_t & = & d_1 \ e[\text{LogTPriceVar}_{t\_}] \\
\theta_t^\alpha & = & d_1^\alpha \ e[\text{LogBuyVolume}_{t\_}] + d_2^\alpha \ e[\text{LogSellVolume}_{t\_}] \\
 & & +d_3^\alpha \ \left| e[\text{LogBuyVolume}_{t\_}] - e[\text{LogSellVolume}_{t\_}] \right| + d_4^\alpha \ E[\text{LogAskDepth}_{t\_}] \\
\theta_t^\beta & = & d_1^\beta \ e[\text{LogBuyVolume}_{t\_}] + d_2^\beta \ e[\text{LogSellVolume}_{t\_}] \\
 & & +d_3^\beta \ \left| e[\text{LogBuyVolume}_{t\_}] - e[\text{LogSellVolume}_{t\_}] \right| + d_4^\beta \ E[\text{LogBidDepth}_{t\_}] \\
\\
\ln(m_t) & = & \ln(m_{t-1}) + \delta_t + \sigma_t \epsilon_t \\
\delta_t & = & c_1^m \ e[\text{LogBuyVolume}_{t\_}] + c_2^m \ e[\text{LogSellVolume}_{t\_}] \\
\epsilon_t & \sim & GED(\nu^{night}) \mathbf{1}_{\{\tau_t^{open}=0\}} + GED(\nu^{day}) \mathbf{1}_{\{\tau_t^{open}>0\}} \\
\ln(\sigma_t^2) & = & \eta_t + \zeta_t + \psi[\ln(\sigma_{t-1}^2) - \eta_{t-1}] + \omega \epsilon_{t-1} + \gamma[|\epsilon_{t-1}| - E|\epsilon_{t-1}|] \\
\eta_t & = & \eta^{night} \mathbf{1}_{\{\tau_t^{open}=0\}} + l_1 + l_2^{open} \exp\left(-l_3^{open} \tau_t^{open}\right) + l_2^{close} \exp\left(-l_3^{close} \tau_t^{close}\right) \\
\zeta_t & = & c_1 \ e[\text{LogSpread}_{t\_}] + c_2 \ e[\text{LogBuyVolume}_{t\_}] + c_3 e[\text{LogSellVolume}_{t\_}] \\
 & & +c_4 \ e[\text{LogSpread}_{(t-1)\_}] + c_5 \ e[\text{LogBuyVolume}_{(t-1)\_}] + c_6 \ e[\text{LogSellVolume}_{(t-1)\_}] \\
 & & +c_7 \ E[\text{LogBidDepth}_{t\_}] + c_8 \ E[\text{LogAskDepth}_{t\_}]
\end{array}
\right.
$$

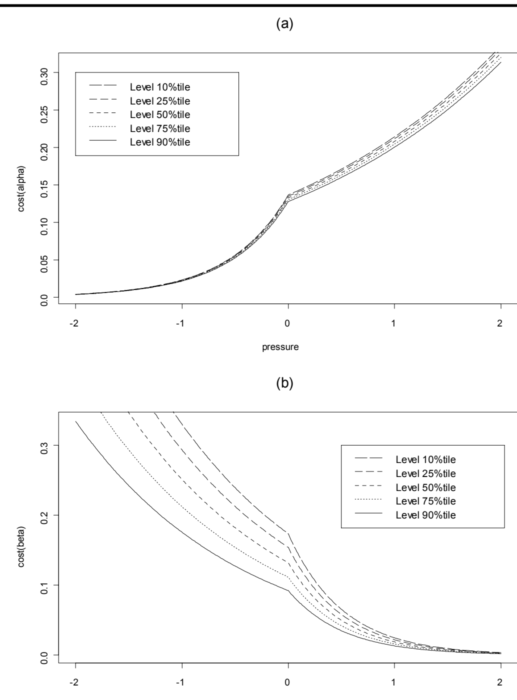
Final Model estimates

Table V: **Parameter Estimates of the Extended Hasbrouck Model for GE**

Cost Model			Efficient Price/EGARCH Model		
Parameter	Posterior Mean	Standard Deviation	Parameter	Posterior Mean	Standard Deviation
ϕ	0.166	0.00835	$c_1^m(e[\text{LogBuyVolume}_{t-}])$	0.000471	0.0000989
σ_v	0.614	0.0561	$c_2^m(e[\text{LogSellVolume}_{t-}])$	-0.000296	0.000127
k_1	-2.05	0.123	ν^{day}	1.59	0.109
k_2^{open}	-0.336	0.222	ν^{night}	2.62	0.675
k_3^{open}	2.31	1.65	ψ	0.86	0.0351
k_2^{close}	0.247	0.177	ω	-0.0169	0.0258
k_3^{close}	4.12	2.23	γ	0.107	0.042
$d_1(e[\text{LogTPriceVar}_{t-}])$	0.201	0.0511	η^{night}	0.495	0.482
$d_1^\alpha(e[\text{LogBuyVolume}_{t-}])$	1.07	0.316	l_1	-12.7	0.0959
$d_2^\alpha(e[\text{LogSellVolume}_{t-}])$	-1.14	0.368	l_2^{open}	2.1	0.428
$d_3^\alpha(\lvert e[\text{LogBuyVolume}_{t-}] - e[\text{LogSellVolume}_{t-}]\rvert)$	-0.661	0.333	l_3^{open}	1.96	0.637
$d_4^\alpha(E[\text{LogAskDepth}_{t-}])$	-0.25	0.155	l_2^{close}	1.21	0.225
$d_1^\beta(e[\text{LogBuyVolume}_{t-}])$	-1.68	0.37	l_3^{close}	3.17	1.3
$d_2^\beta(e[\text{LogSellVolume}_{t-}])$	0.928	0.388	$c_1(e[\text{LogSpread}_{t-}])$	4.23	0.32
$d_3^\beta(\lvert e[\text{LogBuyVolume}_{t-}] - e[\text{LogSellVolume}_{t-}]\rvert)$	-0.656	0.374	$c_2(e[\text{LogBuyVolume}_{t-}])$	0.568	0.0785
$d_4^\beta(E[\text{LogBidDepth}_{t-}])$	-0.487	0.184	$c_3(e[\text{LogSellVolume}_{t-}])$	0.836	0.104
			$c_4(e[\text{LogSpread}_{(t-1)-}])$	-2.92	0.356
			$c_5(e[\text{LogBuyVolume}_{(t-1)-}])$	-0.309	0.0851
			$c_6(e[\text{LogSellVolume}_{(t-1)-}])$	-0.59	0.106
			$c_7(E[\text{LogBidDepth}_{t-}])$	-0.0531	0.0484
			$c_8(E[\text{LogAskDepth}_{t-}])$	-0.094	0.04



More on Volume



Order imbalance interpretations

- Unexpectedly large buyer initiated volume leads to an increase in the ask cost and a decrease in the bid cost.
- Unexpectedly large seller initiated volume leads to an increase in the bid cost and a decrease in the ask cost. The increase in the bid cost is more than twice that of the increase in the ask cost above.

Short sale constraints and volume

- Short sales constraints say that an asset cannot be shorted following a downtick.
- Suppose there is bad news about the asset that a handful of traders know about.
- After the first informed agent sells no other agent can short the stock to capitalize on their better information.
- Hence sell volume may represent only a fraction of the volume that traders would like to have transacted if there were no short sales constraints.
- Conclusion: each unit of seller initiated volume should have a larger impact than each unit of buyer initiated volume.

Some Caveats

- We have interpreted the results as the variables influencing the cost of exposure.
- In fact it could be the other way around.
  - Low cost could induce traders and hence increase overall volume.
  - A movement in the efficient price could induce volume pressure.

Conclusions – Cost Function

- High volatility increases both the cost of purchasing and selling shares.
- Volume effects on the cost are more complex.
  - Unexpectedly high volume generally is associated with lower cost.
  - Imbalance in unexpected volume has asymmetric affects on the cost.
    - Excess buyer volume increases the cost on the ask side and decreases the cost on the bid side.
    - Excess seller volume increases the cost on the bid side **a lot** and decreases the cost on the ask side.

Conclusions Efficient Price

- Unexpectedly large spreads are associated with higher volatility.
- Unexpected volume on either the buy side or the sell side increases the volatility. Unexpected seller initiated volume tends to have a larger immediate impact on volatility.
- Unexpected buyer or seller initiated volume changes the efficient price in the way expected.