F500 Empirical Finance
Lecture 3: Empirical Market Microstructure

Oliver Linton
obl20@cam.ac.uk

January 28, 2020

Outline

1. Microstructure Explanations/Models for serial correlation in stock returns
    1. Stale Prices or infrequent trading. Stale prices are due to infrequent trading (now one usually talks about stale quotes)
    2. Price discreteness: prices are quoted in multiples of the minimum price increment (tick size)
    3. Bid/Ask Bounce
2. Explanations for the existence and magnitude of Bid/Ask (Offer) spread
3. Strategic Trading and the measurement of liquidity
4. Recent Developments

Reading: Linton (2019, Chapter 5), Campbell, Lo and MacKinlay (1997, Chapter 3)

Stale Prices

- CLM model of non trading. There is underlying true price, but it is only observed when a transaction occurs.
- In practice, stocks trade with different frequency, from Apple at one end (many times a millisecond) to "penny stocks" that may only trade once a week. Many empirical questions are concerned with the cross-section of returns, and nonsynchronous trading is a big problem in high frequency data.

- Suppose that the "underlying" price sequence P_t and hence returns r_t for each $t = 1, 2, \ldots$.

$$r_t \text{ i.i.d. } (\mu, \sigma^2)$$

- Suppose that every period

$$\delta_t = \begin{cases} 1 \text{ (no trade)} & \text{with probability } \pi \\ 0 \text{ (trade)} & \text{with probability } 1 - \pi \end{cases}$$

The frequency of trading is controlled by the parameter π. Calendar time model (from lecture 1) is special case with deterministic δ_t.

- Observed prices only update when a transaction takes place

- Observed price p_t^O

$$p_t^O = \begin{cases} p_t & \text{if there is a trade at } t \\ p_{t-1}^O & \text{otherwise} \end{cases}$$

- If $\delta_t = 1$, then

$$r_t^O = p_t^O - p_{t-1}^O = p_{t-1}^O - p_{t-1}^O = 0.$$

- If $\delta_t = 0$, then

$$r_t^O = p_t - p_{t-1}^O,$$

  where p_{t-1}^O may depend on $\delta_{t-1}, \delta_{t-2}, \ldots$
- The key quantity is how stale is the price p_{t-1}^O, which depends on how many periods have $\delta_{t-j} = 1$

Definition

The duration of non trading, denoted d_t, is given by the integer k for which $\delta_{t-k} = \delta_t = 0$ but $\delta_s = 1$ for $s \in (t-k, t)$ and $d_t = 0$ if t and $t-1$ have trades.

- The random variable $d_t \in \{0, 1, 2, \ldots\}$, and depends on the past sequence of $\{\delta_t\}$.
- It has a dynamic evolution

$$d_{t+1} = \begin{cases} d_t + 1 & \text{with probability } \pi \\ 0 & \text{with probability } 1 - \pi. \end{cases}$$

This is a stationary Markov stochastic process that evolves over time.

The observed (logarithmic) return is $r_t^O = p_t^O - p_{t-1}^O$.

$$r_t^O = \begin{cases} 0 & \text{with prob } \pi \text{ (no trade today)} \\ r_t & \text{with prob } (1-\pi)^2 \text{ (trade today and yday)} \\ r_t + r_{t-1} & \text{with prob } (1-\pi)^2\pi \text{ (trade today and day b4 yday)} \\ \vdots & \vdots \end{cases}$$

$$r_t^O = \begin{cases} 0 & \text{with prob } \pi \\ \sum_{k=0}^{d_t} r_{t-k} & \text{with prob } 1-\pi \end{cases}$$

true price	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}
no trade? δ	0	0	0	1	1	0	0	0	1	1
observed price	p_1	p_2	p_3	p_3	p_3	p_6	p_7	p_8	p_8	p_8
true return		r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}
obs return		r_2	r_3	0	0	$r_4 + r_5 + r_6$	r_7	r_8	0	0
duration d	0	0	0	1	2	0	0	0	1	2

The marginal distribution of d_t is Geometric (type 2) with $p = 1 - \pi$. That is,

$$d_t = \begin{cases} 0 & 1 - \pi \\ 1 & (1 - \pi)\pi \\ 2 & (1 - \pi)\pi^2 \\ \vdots & \vdots \end{cases}$$

Therefore, can show

$$Ed_t = \frac{\pi}{1 - \pi}$$

$$\mathrm{var}(d_t) = \frac{\pi}{(1 - \pi)^2}$$

Main Implications for single stock

- We can show that (for $n = 1, 2, \ldots$):

$$\begin{aligned} E(r_t^O) &= \mu \\ \mathrm{var}(r_t^O) &= \sigma^2 + \frac{2\pi}{1-\pi}\mu^2 \\ \mathrm{cov}(r_t^O, r_{t+n}^O) &= -\mu^2\pi^n. \end{aligned}$$

- This is consistent with observed returns following an ARMA(1,1) process such that

$$r_t^O - \mu = \pi(r_{t-1}^O - \mu) + \eta_t + \theta\eta_{t-1},$$

where η_t is iid mean zero with variance $\sigma_\eta^2(\pi, \mu, \sigma^2)$ and $\theta(\pi, \mu, \sigma^2)$ is such that $|\theta| < 1$.

Cross Covariances

- We consider the bivariate case with

$$\mathrm{cov}(r_{it}, r_{jt}) = \sigma_{ij}.$$

  We find empirically that $\sigma_{ij} >> 0$ for most pairs of stocks. Each security has non trading probability π_i.
- Can show that for $i \neq j$, $n \geq 0$

$$\gamma_{ij}(n) = \mathrm{cov}(r_{it}^{O}, r_{jt+n}^{O}) = \frac{(1-\pi_i)(1-\pi_j)}{1-\pi_i\pi_j}\sigma_{ij}\ \pi_j^n$$

- Predicts that cross-autocorrelations are the same sign as σ_{ij}, ie usually positive. Note $\frac{(1-\pi_i)(1-\pi_j)}{1-\pi_i\pi_j} \leq 1$ so $\mathrm{cov}(r_{it}^{O}, r_{jt}^{O}) \leq \mathrm{cov}(r_{it}, r_{jt})$.
- These formulae explain some of patterns in autocorrelations and cross autocorrelations in individual stocks reported in CLM

- Relative cross-covariances depend only upon the trading frequencies of the securities

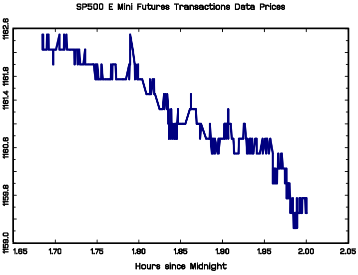
$$\frac{\gamma_{ij}(n)}{\gamma_{ji}(n)} = \frac{\mathrm{cov}(r_{it}^{O}, r_{jt+n}^{O})}{\mathrm{cov}(r_{jt}^{O}, r_{it+n}^{O})} = \left(\frac{\pi_j}{\pi_i}\right)^n.$$

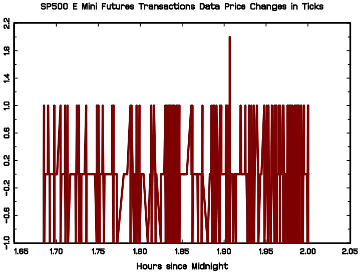
- Note that the more liquid stock (small π) tends to lead more strongly and the less liquid stock tends to lag more strongly, but both lead and lag effects are present unless $\pi = 1$ for one stock.

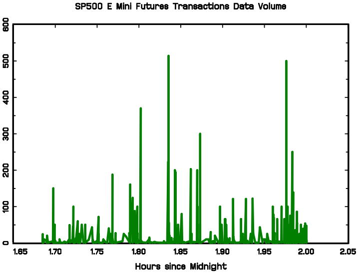
- The non trading model gives some useful insights, especially with regard to non synchronous trading. However:
  - CLM model is nominally about daily data (in 1980s this made sense) Nowadays all S&P500 stocks, say, trade every day many times now, so perhaps more relevant for small stocks or corporate bonds or intraday data
  - Problem for intraday application is have to choose some unit of time: In a trading day (LSE is 0800-1630 and NYSE is 0930-1600) there may be 400 minutes, 25000 seconds, 25000000 milliseconds etc.]
  - Observable quoted prices contain information beyond that contained in the most recent trade and are updated more frequently than trades. Use the midquote.
  - Magnitudes of effects too small to explain autocorrelation properties (CLM)
  - Only relevant now for intraday transactions data. In that case, π is surely not fixed within a day and depends on past trades and prices as well as time of day. Engle ACD model and many others push this forward.

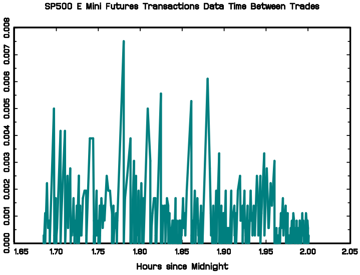
Discreteness

- Quantity and Prices are discrete (minimum price increment, minimum quantity). In 1997 when CLM was published, the minimum price increment or tick size in the US was 1/8th of a dollar
- In the United States, for any stock over \$1 in price level the tick size is currently one cent (although there are exceptions - Berkshire Hathaway A priced at \$134,060.00 only takes \$1 moves apparently, although http://www.nanex.net/aqck2/3571.html).
- US is currently debating "subpenny pricing" whereby tick size might be reduced to 0.1 of a cent. In FX, tick size can be 0.0001 or smaller.
- In UK (and most countries except USA), tick size varies across stocks in bands according to the price level and market capitalization (generally speaking more liquid stocks have small tick sizes). Tick size has been subject to regulatory debate.


SP500 E Mini Futures Transactions Data Prices
1159.0
1159.8
1160.6
1161.4
1161.8
1162.6
1.65
1.70
1.75
1.80
1.85
1.90
1.95
2.00
2.05
Hours since Midnight


SP500 E Mini Futures Transactions Data Price Changes in Ticks
-1.0
-0.6
-0.2
0.2
0.6
1.0
1.4
1.8
2.2
1.65
1.70
1.75
1.80
1.85
1.90
1.95
2.00
2.05
Hours since Midnight


SP500 E Mini Futures Transactions Data Volume
0
100
200
300
400
500
600
1.65
1.70
1.75
1.80
1.85
1.90
1.95
2.00
2.05
Hours since Midnight


SP500 E Mini Futures Transactions Data Time Between Trades
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
1.65
1.70
1.75
1.80
1.85
1.90
1.95
2.00
2.05
Hours since Midnight

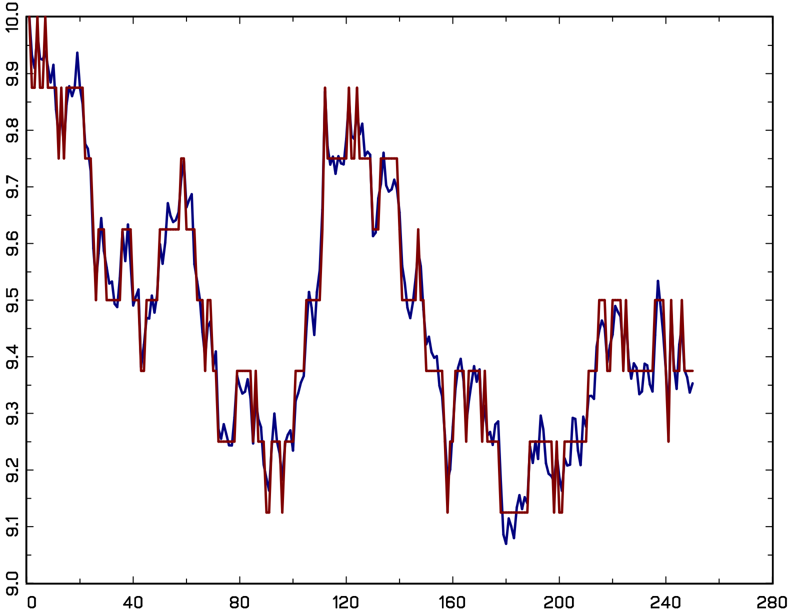
- Consequence of discreteness is negative autocorrelation in observed price changes
- Why? The logic is a bit like a non trading model except it also works even if there is no drift. If the rounding bites often, then many returns are zero and positive return is most likely followed by zero return and negative return is most likely followed by zero return.
- Suppose that

$$P_t = P_{t-1} + \varepsilon_t$$

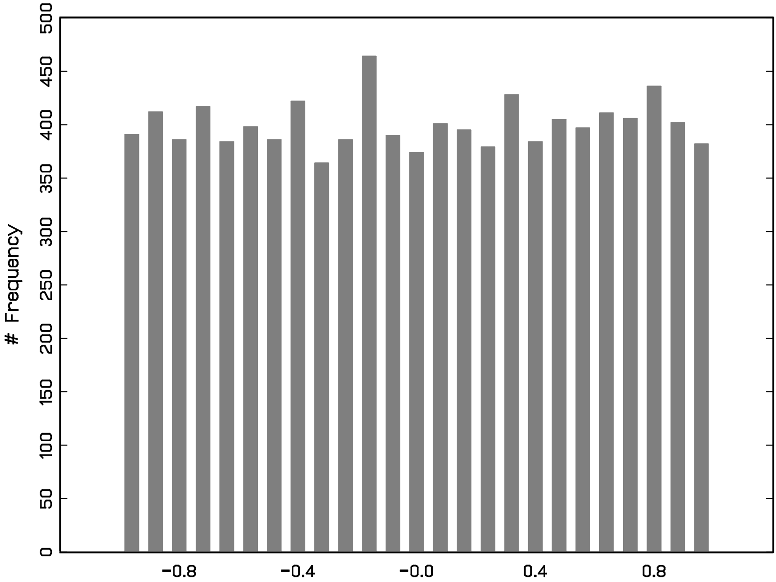
where ε_t is $N(0, \sigma^2)$, $P_0 = 10$ and $\sigma = 0.1$, and let

$$P_t^R = round(P_t, 1/8)$$

which means the closest point to P_t on the grid $\{j/8, j = 1, 2, \ldots\}$. Then $P_t^R - P_{t-1}^R$ is negatively autocorrelated at first lag.


10.0
9.9
9.8
9.7
9.6
9.5
9.4
9.3
9.2
9.1
9.0
0
40
80
120
160
200
240
280

Round off errors are approximately uniform. We show histogram of $16 * (P_t - P_t^R)$



Bid, Ask and Transaction Prices

Dealer market: Dealer/market maker quote bid and ask prices either as needed or displayed publicly (think Travelex). Take it or leave it. He knows the flow of orders. In some case is a monopolist and has unique access to the order flow information; in other cases competitive dealers. Nowadays most equity markets are organized as Electronic Order Books Anyone can enter buy or sell orders. Transparent display of demand and supply. Competition between liquidity providers.

The theoretical literature is mostly concerned with a stylized version of the dealer market where:

- Market participants can be sorted into outsiders (everyone except the market maker) and insiders (the market maker).
- The market maker sets the ask price and the bid price. He hopes to earn the bid ask spread, i.e., to buy at the bid and sell at the ask.
- If she can make these two trades at exactly the same time, this is a money printing machine. But she cant and it is generally a risky business as we see.

Roll Model

- Roll model: What is the consequence of a (fixed) bid-ask spread for observed transaction price changes?
- Assume that fundamental price P^* is a random walk

$$P_t^* = P_{t-1}^* + \varepsilon_t.$$

- Dealer sets ask price $P_t = P_t^* + \frac{s}{2}$ and bid price $P_t = P_t^* - \frac{s}{2}$, where s is the spread. Buy and sell orders arrive randomly unrelated to (independent of) fundamental price. Transaction price is

$$P_t = P_t^* + Q_t \frac{s}{2},$$

  where Q_t is the customer (market) order direction indicator, $+1$ for buy and -1 if customer is selling. A special case of the Fads model!
- It follows that

$$\Delta P_t = \Delta P_t^* + (Q_t - Q_{t-1}) \frac{s}{2}$$

- Assume that Q_t is iid with equal probability of $+1$ and -1 and unrelated to P_t^*.
- Since by definition of a random walk, P^* has zero autocovariance and by Roll's assumption Q_t is unrelated to P^*

$$\mathrm{cov}\left[\Delta P_{t-1}, \Delta P_t\right] = -\frac{s^2}{4}$$

  This is called *Bid Ask Bounce* (BAB) - presence of bid ask spread induces negative first order autocorrelation in transaction price.
- Note that ΔP is an MA(1) process with

$$\mathrm{cov}\left[\Delta P_{t-j}, \Delta P_t\right] = 0,\ j \geq 2$$

- Multivariate case with cross-correlated order flow and heterogeneous spreads leads to Vector MA(1) process
- The midquote $M_t = P_t^*$ is serially uncorrelated.

- One application of this is to give an estimating equation for the spread, s, (which is the cost to traders of trading)

$$s = 2\sqrt{-\operatorname{cov}\left[\Delta P_{t-1}, \Delta P_t\right]}$$

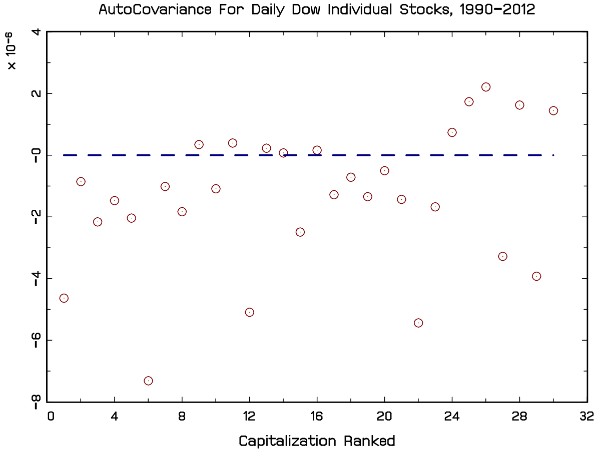
This is sometimes used to estimate the spread (when this is not directly observed) from transaction data. If use returns instead of price changes get percentage spread.

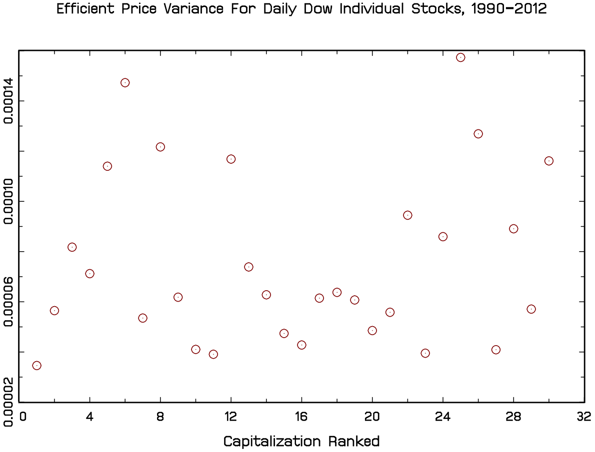
- The noise of the random walk (fundamental volatility) is

$$\sigma_\varepsilon^2 = \operatorname{var}\left[\Delta P_{t-1}\right] + 2\operatorname{cov}\left[\Delta P_{t-1}, \Delta P_t\right].$$

This is of interest to compare with total volatility

- Empirically problematic. Many autocovariances are positive. A number of refinements to this measure that address this issue have been suggested, see for example Hasbrouck (2006).


AutoCovariance For Daily Dow Individual Stocks, 1990-2012
x 10^-6
4
2
-0
-2
-4
-6
-8
0
4
8
12
16
20
24
28
32
Capitalization Ranked


Efficient Price Variance For Daily Dow Individual Stocks, 1990-2012
0.00002
0.00006
0.00010
0.00014
0
4
8
12
16
20
24
28
32
Capitalization Ranked

Market Microstructure Models

- What determines the bid-ask spread?
  - Inventory models: The bid ask spread compensates the market maker for the risk of ruin through inventory explosion
  - Information models. Bid-ask spread is determined by adverse selection costs
  - Combinations allow both adverse selection and inventory costs.

Inventory Models

- Important role of market makers: provide opportunity to trade at all times ("immediacy") to traders
- Market makers (MM) absorb temporary imbalances in order flow and will hold inventory of assets. Inventory may deviate from desired inventory position (in the long run, zero) due to changes in demand and supply
- Market maker requires compensation for service of providing "immediacy". Ho and Stoll (1981,1983). Solves a dynamic optimization problem in continuous time.
- The bid ask spread S is independent of the current inventory level and

$$S \simeq A + \sigma^2 \theta T q$$

  where: A is a fixed measure of the monopoly power of the MM, q is the trade size for which the quote is relevant, σ^2 is volatility of stock price, θ is a measure of the risk aversion of the MM, and T is his (short) time horizon

Information Models

- Glosten-Milgrom (1985) models. Asymmetric information. Informed traders and uninformed traders.
- Adverse selection is an important issue for market makers/dealers. Harder to price discriminate as insurance companies do to mitigate risks.
- The bid-ask spread is compensation for the market maker's adverse selection costs.

A simple sequential trade model

Glosten-Milgrom (1985). Hasbrouck (2007, pp44-46).

- Value V is chosen from the distribution

$$V = \begin{cases} V_H & \text{with prob } 1-\delta \\ V_L & \text{with prob } \delta \end{cases}$$

- Type of investor is chosen each period from

$$T = \begin{cases} I & \text{with prob } \mu \\ U & \text{with prob } 1-\mu \end{cases}$$

- Strategies:
  - If informed (I), buy if value is high V_H and sell if value is low V_L
  - If uninformed (U), buy or sell with probability $1/2$
  - Dealer sets B, A to make zero profits (this is forced on her by competition or regulation). Observes noisy signal, the order flow

- Each period there is a buy or sell order. Let Q be the indicator of order direction, i.e., $Q = +1$ if the order is a buy order and $Q = -1$ if the order is a sell order.
- The order flow is described by the following probability table:

	$Q = +1$	$Q = +1$
$V = V_L$	$\frac{1}{2}(1-\mu)$	$\frac{1}{2}(1+\mu)$
$V = V_H$	$\frac{1}{2}(1+\mu)$	$\frac{1}{2}(1-\mu)$

where for example the probability of receiving a buy order when the asset has low value is $\Pr(Q = +1|V = V_L) = (1-\mu)/2$, just coming from uninformed traders.

- Dealer reasons: if receive a buy order how would I update my valuation?
- By Bayes rule, she would calculate posterior distribution given Q

$$\overbrace{\Pr(V = V_L | Q = +1)}^{\textit{posterior}} = \frac{\overbrace{\Pr(Q = +1 | V = V_L)}^{\textit{likelihood}} \overbrace{\Pr(V = V_L)}^{\textit{prior}}}{\Pr(Q = +1)}$$

$$= \frac{\frac{1}{2}(1-\mu) \times \delta}{(1 + \mu(1 - 2\delta))/2}$$

- Then $\Pr(V = V_H | Q = +1) = 1 - \Pr(V = V_L | Q = +1)$. Likewise compute $\Pr(V = V_L | Q = -1)$ and $\Pr(V = V_H | Q = -1)$ and use to calculate $E(V | Q = \pm 1)$

Zero expected profit condition (side by side) implies that:

$$A = E(V|Q=+1) = V_L \frac{\delta(1-\mu)}{1+\mu(1-2\delta)} + V_H \frac{(1-\delta)(1+\mu)}{1+\mu(1-2\delta)}$$
$$B = E(V|Q=-1) = \frac{V_L \delta(1+\mu) + V_H(1-\delta)(1-\mu)}{1-\mu(1-2\delta)}$$

Therefore, bid-ask spread is

$$s = A - B = \frac{4(1-\delta)\delta\mu}{1-(1-2\delta)^2\mu^2}(V_H - V_L)$$

- When $\delta = 1/2,$ the bid-ask spread is

$$s = A - B = \mu(V_H - V_L)$$

- Bid-ask spread is wider when
  - There are more informed investors
  - When $V_H - V_L$ is larger (could interpret this as volatility or uncertainty over final value)
- Dealer gains from uninformed traders and loses to informed ones.
- The signal from the order direction adjusts the mean and reduces the uncertainty about the value

$$E(V|Q = +1) = \frac{1}{2} V_L (1 - \mu) + \frac{1}{2} V_H (1 + \mu) = E(V) + \frac{\mu}{2} (V_H - V_L)$$

$$\mathrm{var}(V|Q = +1) = \frac{1}{4} (1 - \mu^2) (V_H - V_L)^2 \leq \frac{1}{4} (V_H - V_L)^2 = \mathrm{var}(V)$$

- Midpoint is $M = (V_L + V_H)/2 = E(V)$ the unconditional expectation of value.

- Now consider this process evolving over time (with new traders every period but constant values V_H, V_L).
- Dealer observes a price/order history at any time $t-1$, for example $\{Q_1 = 1, Q_2 = -1, Q_3 = -1, \ldots, Q_{t-1} = 1\}$, denote this by $\mathcal{F}_{t-1}$. Orders are serially correlated - informed investors always trade in the same direction. He updates his valuation of the security given $\mathcal{F}_{t-1}$.
- Prior is now replaced by the posterior distribution from the last round $\Pr(V = V_L | \mathcal{F}_{t-1})$, whence obtain the posterior

$$\Pr(V = V_L | Q_t = 1, \mathcal{F}_{t-1}) = \frac{\Pr(Q_t = 1 | V = V_L, \mathcal{F}_{t-1}) \Pr(V = V_L | \mathcal{F}_{t-}}{\Pr(Q_t = 1 | \mathcal{F}_{t-1})}$$

- Sets the next period ask and bid price by the same zero profit condition

$$A_t = E(V | Q_t = 1, \mathcal{F}_{t-1}) \quad ; \quad B_t = E(V | Q_t = -1, \mathcal{F}_{t-1})$$

- The transaction price is determined by what the actual incoming order is

$$P_t = \begin{cases} A_t & \text{if } Q_t = +1 \\ B_t & \text{if } Q_t = -1 \end{cases} = M_t + Q_t \frac{s_t}{2}$$

- This is a martingale with respect to the upto date history $\mathcal{F}_t$ because $P_t = E(V|\mathcal{F}_t)$ and so

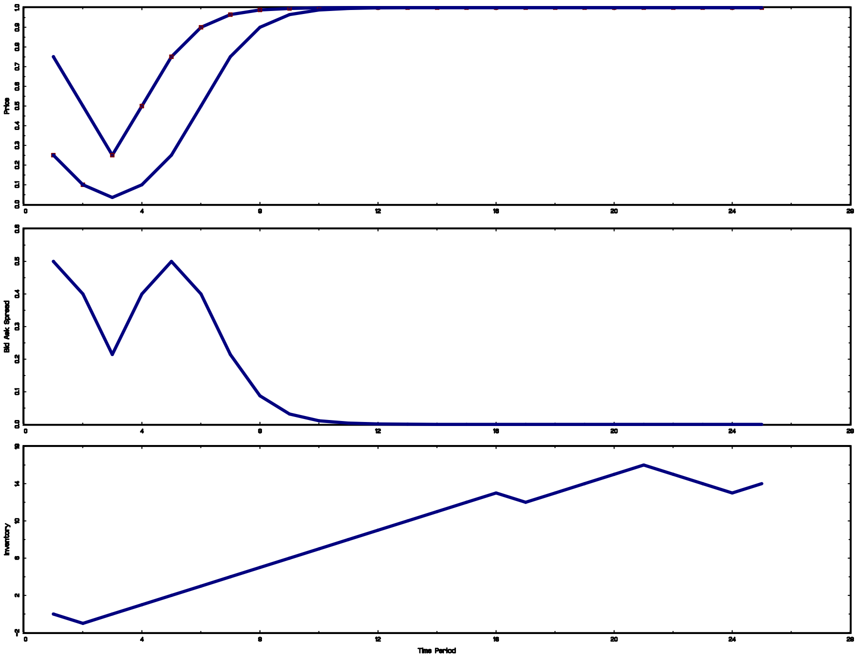
$$E\left[P_{t+1}|\mathcal{F}_t\right] = E\left[E\left[V|\mathcal{F}_{t+1}\right]|\mathcal{F}_t\right] = E\left[V|\mathcal{F}_t\right] = P_t$$

  by the law of iterated expectation, $\mathcal{F}_t \subset \mathcal{F}_{t+1}$.
- This says that microstructure effects do not necessarily lead to serial correlation.

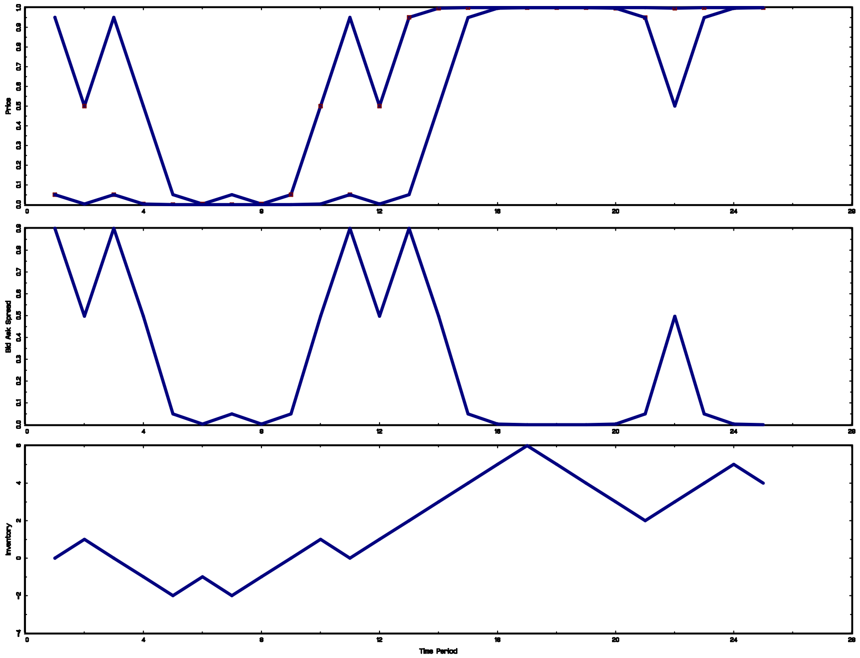
- Information is improving over time:
  - Trades (or rather order flow) have price impact because they lead to an adjustment in the posterior
  - Market maker can consistently estimate the true value from a long sequence of orders
  - Spreads get narrower over time
  - Prices converge to the true value. Eventually, all the informed trader's information is absorbed in price

Simulated examples. Chose $V = V_H = 1$ with $V_L = 0$
1. $\mu = 0.5, \delta = 0.5$
2. $\mu = 0.1, \delta = 0.5$
Shows as time goes by

- in panel A, Ask price, Bid price, transaction price (red squares)
- in panel B, bid ask spread
- in panel C, inventory of dealer



Price
Bid Ask Spread
Inventory
Time Period


Price
Bid Ask Spread
Inventory
Time Period

Adverse Selection plus Inventory Costs

- Glosten (1987). Considers more general model where bid ask spread is determined by a combination of: adverse selection, order processing costs, inventory costs, and monopoly power. In this case, transaction price is no longer a martingale (stock returns are serially correlated):

$$P_{t+1} = E(V|\mathcal{F}_t) + \overbrace{Q_{t+1}c}^{\text{fixed cost}}$$

- Order processing costs etc are temporary in effect, whereas information has permanent effect on prices.
- Glosten and Harris (1988). Develop empirical model containing adverse selection and inventory components. They find that the adverse selection component is significant and is related to relevant stock characteristics.

- More realistic models (Easley and O'Hara (1987, 1992) (VPIN) allows
  - Informed investors do not know precisely the value of stock but observe private signals (information events).
  - Not all trading days have information events
  - Trades can be of different sizes and price impact depends on trade size
  - Participants may or may not trade so the time between trades is variable and this may have information content

Strategic trade models

- Kyle model (1985). Single informed trader who knows their own impact on prices and trades strategically, i.e., not all or nothing. Continuous time version of model: informed trader chops up their order and feeds into the market. We consider just single period version.
- Value of security $v \sim N(\mu_v, \sigma_v^2)$.
  - Informed trader knows v and submits demand $x(v)$ to maximize his expected profit
  - Noise traders submit order flow $u \sim N(0, \sigma_u^2)$
  - Risk neutral Market maker observes total demand $y = x + u$ and then sets a price p to make zero profits in expectation
- Market maker uses Bayes theorem to try to estimate $E(x|y)$

- Equilibrium is a pricing rule of the market maker $p(y)$ and informed trader demand function $x(v)$ that are mutually consistent.
- Can show that such equilibrium exists and that

$$p(y) = \mu_v + \frac{1}{2}\sqrt{\frac{\sigma_v^2}{\sigma_u^2}}y$$

$$x(v) = -\mu_v\sqrt{\frac{\sigma_u^2}{\sigma_v^2}} + \sqrt{\frac{\sigma_u^2}{\sigma_v^2}}v$$

- The informed trader makes expected profit

$$\frac{(v-\mu_v)^2}{2}\sqrt{\frac{\sigma_u^2}{\sigma_v^2}} \geq 0,$$

  which is at the expense of the noise traders. MM breaks even in expectation.
- Half of the informed trader's information is impounded in price $\mathrm{var}[v|y] = \sigma_v^2/2$

Definition

Kyle's Lambda

$$p(y) = \mu_v + \lambda y, \quad \lambda = \frac{1}{2}\sqrt{\frac{\sigma_v^2}{\sigma_u^2}}$$

- This is the amount that the market maker raises the price when the total order flow y goes up by 1 unit ($\lambda = dp/dy$). Hence, the amount of order flow necessary to raise the price by \$1 equals $1/\lambda$, which is a measure of the "depth" of the market or market "liquidity."
- The higher is the proportion of noise trading to the value of insider information, the deeper or more liquid is the market. Intuitively, the more noise traders relative to the value of insider information, the less the market maker needs to adjust the price in response to a given order, since the likelihood of the order being that of a noise trader, rather than an insider, is greater.

- There are a number of ways of measuring liquidity from low frequency data, see Goyenko et. al. (2009, JFE) for a review. Amihud (2002) develops a liquidity measure that captures the daily price response associated with one unit of trading volume.

Definition

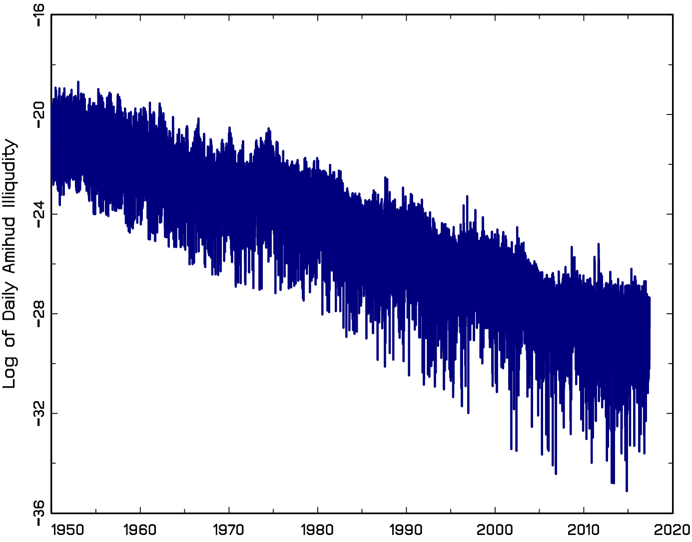
Let

$$Am_t = \frac{1}{N_t}\sum_{j=1}^{N_t} \ell_{t_j}, \quad \ell_{t_j} = \frac{|R_{t_j}|}{V_{t_j}},$$

where V is the trading volume and R returns. Here, we should average daily ℓ_{t_j} over a longer period like a week or a month.

- Large values of this measure indicate an illiquid market where small amounts of volume can generate big price moves. It is considered a good proxy for the theoretically founded Kyle's price impact coefficient.

- This shows the raw unaveraged ℓ_{t_j} for daily data



Daily Amihud illiquidity on S&P500

The limit order book

Bid		Ask	
Price	Quantity	Price	Quantity
15.71	2000	15.72	7000
15.70	4500	15.73	3000
15.69	5000	15.74	4000
15.68	10000	15.75	2000
15.67	15000	15.76	12000

At any one time, this is available to (some) participants.
Market orders or aggressive limit orders (that cross the spread) will execute against "the book".
Market buy order for 15000: 7000@15.72, 3000@15.73, 4000@15.74@, 1000@15.75 (walking down the book). Volume weighted average price (VWAP)

$$VWAP = \frac{7}{15} \times 15.72 + \frac{3}{15} \times 15.73 + \frac{4}{15} \times 15.74 + \frac{1}{15} \times 15.75 = 15.729$$

New order book. Bid ask spread is 0.04.

Bid		Ask	
Price	Quantity	Price	Quantity
15.71	2000	15.75	1000
15.70	4500	15.76	12000
15.69	5000		
15.68	10000		
15.67	15000		

Until replenished by new limit orders

This is (mechanical) **market impact**

If the market buy order were for quantity 2000, then the spread would not change. The order executes by Price Time Priority - the first order at the price gets executed first and so on

Real time limit order book: http://www.batstrading.co.uk/

- Posting limit orders (supplying liquidity) gives options to trade to other traders. Provides a service to other traders. Needs to be compensated.
- The value of those options can be calculated from Black and Scholes (1973, JPE) call option price

$$C(S, K, \tau, r_f, \sigma) = S \cdot \Phi(d_+) - K \cdot e^{-r_f \cdot \tau} \cdot \Phi(d_-)$$

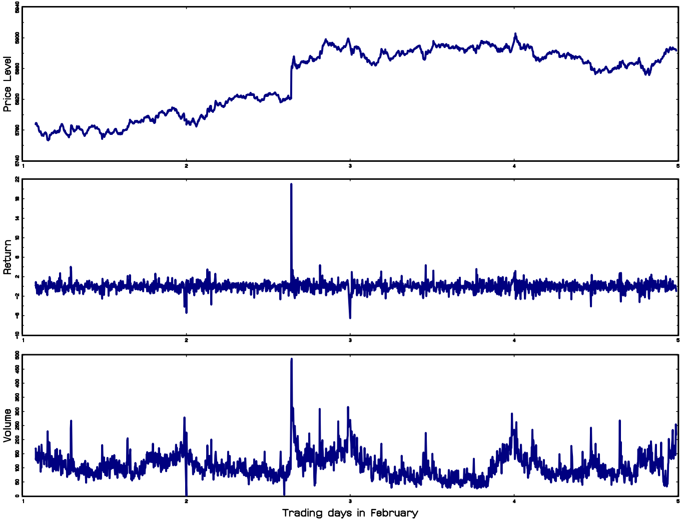
$$d_\pm = \frac{\log \frac{S}{K} + \left(r_f \pm \frac{\sigma^2}{2} \right) \cdot \tau}{\sigma \cdot \sqrt{\tau}}$$

and Φ is the standard normal cdf. At the money, $S = K$, as $\tau \to 0$

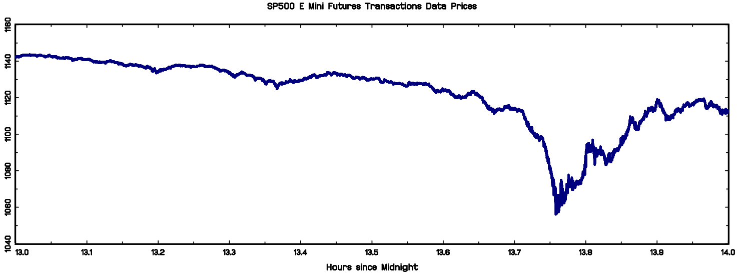
$$C(\overbrace{S}^{\text{spot}}, \overbrace{K}^{\text{strike}}, \overbrace{\tau}^{\text{time to maturity}}, \overbrace{r_f}^{\text{interest}}, \overbrace{\sigma}^{\text{volatility}}) = \frac{1}{\sqrt{2\pi}} S \cdot \sigma \sqrt{\tau} + O(\tau).$$

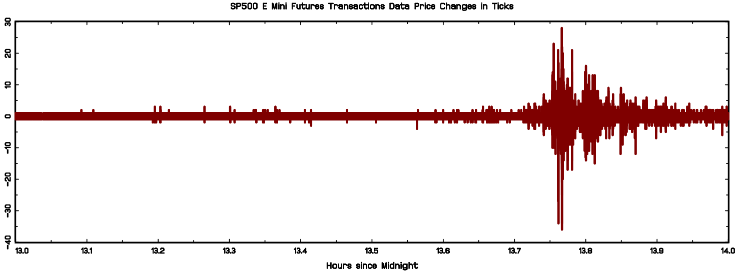
- There is a positive albeit small value in an order that only sits for small time.
- If there are many orders, then value =large$\times$*small*

Markets import new information quickly



Flash Crash in the US stock market on May 6th, 2010. Partly blamed on High Frequency (Computer) Trading (HFT) and Algorithmic (Computer) Trading (AT).





Appendix

Since $\{\delta_t\}$ is independent of $\{r_t\}$ we can condition on the sequence $\{\delta_t\}$ and use the law of iterated expectation

$$
\begin{aligned}
Er_t^O &= (1-\pi)E\left[E\left(\sum_{k=0}^{d_t} r_{t-k}\,|d_t\right)\right] = (1-\pi)\mu E(d_t+1) = \mu \\
E[(r_t^O)^2] &= (1-\pi)E\left[E\left(\left(\sum_{k=0}^{d_t} r_{t-k}\right)^2 |d_t\right)\right] \\
&= (1-\pi)\left[E(d_t+1)Er_t^2 + E((d_t+1)d_t)E^2 r_t\right] \\
&= (1-\pi)\left[E(d_t+1)(\sigma^2+\mu^2) + E((d_t+1)d_t)\mu^2\right] \\
&= (1-\pi)\left[\frac{1}{1-\pi}(\sigma^2+\mu^2) + \frac{2\pi}{(1-\pi)^2}\mu^2\right]
\end{aligned}
$$

We calculate $E[r_t^O r_{t+1}^O]$ by first conditioning on the sequence $\{\delta_t\}$. Unless $\delta_{t+1} = \delta_t = 0$, we have $r_t^O r_{t+1}^O = 0$, so we only need consider the case that $\delta_{t+1} = \delta_t = 0$, which has probability $(1-\pi)^2$. In this case, $d_{t+1} = 0$. It follows that

$$\begin{aligned} E[r_t^O r_{t+1}^O] &= (1-\pi)^2 E\left(r_{t+1} \sum_{i=0}^{d_t} r_{t-i} \right) \\ &= (1-\pi)^2 \mu^2 E(d_t + 1) \\ &= \mu^2 (1-\pi)^2 \left(\frac{\pi}{1-\pi} + 1 \right) \\ &= \mu^2 (1-\pi). \end{aligned}$$

Therefore, as in CLM (1997, p89)

$$\mathrm{cov}(r_t^O, r_{t+1}^O) = -\pi\mu^2.$$