Financial Markets Microstructure
Lecture 8

Empirics of illiquidity
Chapter 5 of FPR

Egor Starkov

Københavns Unversitet
Spring 2025

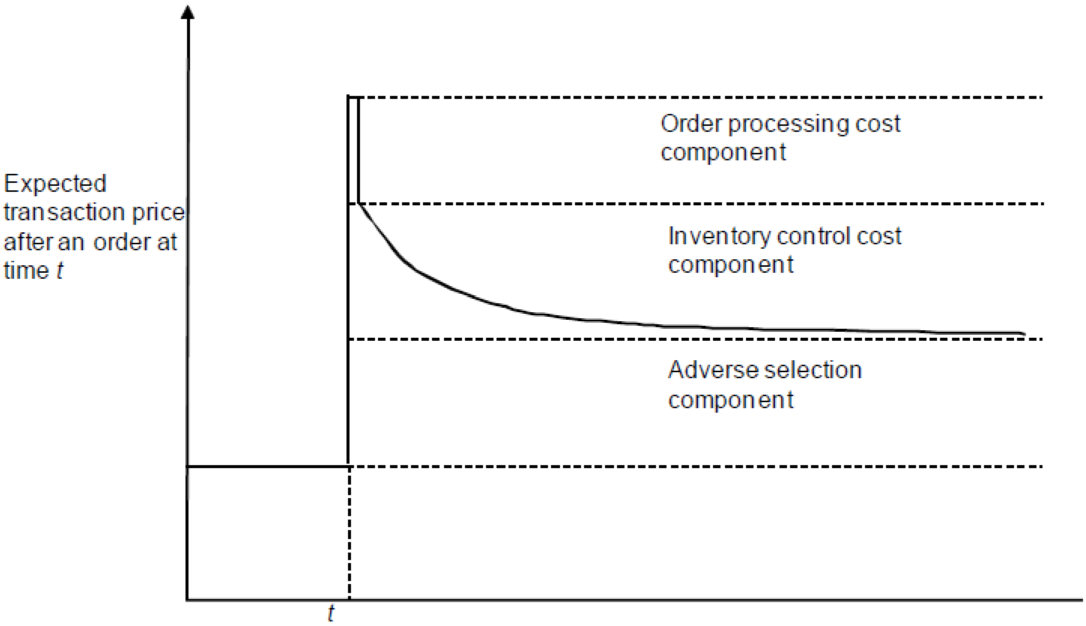
What did we do last week?

1. Depth is driven by mostly same factors as liquidity: adverse selection and inventory risk
   - Although the effect is, in general, less obvious with order costs and imperfectly competitive dealers
2. Kyle model demonstrates how adverse selection can drive limited depth
   - Call auction market where dealer has to cover the order imbalance
   - Informed traders face a trade-off: trade more vs trade at better price
   - see textbook for Kyle model with inventory risk.

Today

Empirical estimation of liquidity

- Earlier we looked at how to estimate the spread, but without a theory for what drives it
- Today: talk about estimating drivers of the spread
  - Knowing what drives illiquidity helps understand how to eliminate (or exploit) it
- Note: most estimates mentioned today are pretty old, and I could not find a lot of more recent ones – a fine thesis topic!



Measured price impact and the theories

- Parameters of interest:
  - λ: price impact related to information
  - β: price impact related to MM risk aversion
  - γ: order-processing cost (including dealer profits!)
- Data:
  - Transaction prices p_t, net market order flow q_t, order sign d_t

How to estimate?

- Consider a market with adverse selection and order costs.
- Suppose prices are ex post efficient (apart from order costs), then:

$$p_t = \mu_t + \gamma d_t$$
$$\Rightarrow \Delta p_t = \mu_t - \mu_{t-1} + \gamma \Delta d_t$$

- and market valuation μ_t evolves due to order flow q_t and other public news ε_t:

$$\mu_t = \mu_{t-1} + \lambda q_t + \varepsilon_t$$

- Then

$$\Delta p_t = \lambda q_t + \gamma \Delta d_t + \varepsilon_t$$

?

- **?** estimate

$$\Delta p_t = \gamma_0 \Delta d_t + \gamma_1 \Delta q_t + \lambda_0 d_t + \lambda_1 q_t + \epsilon_t \tag{5.7}$$

  jointly with trade directions d_t.

- Accumulated order flow affects price level via λ, cyclical order flow gives γ in the style of Roll's measure (2.15)
- On NYSE data from early 80s, they find $\gamma_1 = \lambda_0 = 0$ and estimate $\gamma_0 = .0465$, $\lambda_1 = .0102$ for average stock in sample (so a \$1000 trade pays \$56.7 total in trading costs and permanently shifts the price by \$10.2)

Issues

- Apart from a few ad hoc concerns...
  - not much data (20 stocks, 800 transactions each)
  - poor dataset (neither d_t, nor quotes are observed; prices p_t were rounded to \$1/8 back then)
  - questionable specification (why have both λ_0 and λ_1?)
- ...there is the issue of ignoring inventory costs.
- In the short run, the effect of MM risk aversion would be similar

$$\Delta p_t = \gamma \Delta d_t + (\lambda + \beta) q_t + \epsilon_t \tag{5.14}$$

  - λ and β are not separately identified.

Including inventory costs

- **?**: order flow is auto-correlated, so part of the order is not 'news' but only moves inventory
- **?**: suggest

$$q_t = \phi q_{t-1} + \eta_t \qquad (5.17)$$

- Then (5.14) becomes Derivation

$$\Delta p_t = \gamma \Delta d_t + \lambda(q_t - \phi q_{t-1}) + \beta q_t + \epsilon_t \qquad (5.21)$$

- They estimate the model for 20 major NYSE stocks (mostly ignoring volumes)

$$\phi = -0.74 \qquad \lambda = 9.59\%S \qquad \beta = 28.65\%S \qquad \gamma = 61.76\%S$$

- Also: greater β for larger transactions

Including inventory costs (2)

- **?** find that λ is higher in the morning, and $\beta + \gamma$ higher in the afternoon
  - Morning trading reveals all the information accumulated during off-market hours. Evening trading driven by traders' desire to close open positions
  - **?**: closing auctions account for 7.3% of daily trading volume, but contribute nothing to price discovery. Closing prices deviate a lot from midquotes, but revert back quickly. Why? Probably because traders rush to close their positions before EoD. But also in part because closing auctions are more closed and are of lower quality, see **?**.
- **?** had data on a FOREX dealer's inventory, and found β very large
  - Probably also because FX dealers really don't like to hold overnight positions

Long-run impact

- Different dynamic approach: **?** uses the long-run impact of an order q_{t+1} on prices $p_{t+T} - p_t$ to identify the informational part of the trade
  - The 'impulse response' of prices to orders
  - Opens door to richer time series analysis of transactions data
- Findings: short-run effect of a trade $\approx 2\%$ of the price
- long-run effect of a trade ($T = 5$ trades is "long enough") $\approx 1\%$
  - greater LR impact for stocks with lower capitalization
  - related: **?** find that stocks with lower analyst coverage are less liquid

Depth impact

- **?** do analysis similar to **?** with data from Paris Bourse, and taking trade quantities into account
- Find that LR price impact is between 25% and 115% of the spread (depending on estimation method)
- Find that adverse selection costs increase with trade size...
  - aligns with Kyle model
- ...while the converse is true for order processing costs. Possible reasons:
  - costs are per-transaction rather than per-unit
  - large orders come from institutional traders that get better terms of trade? (However: Paris Bourse was a pure-LOB market, no dealers, so this is probably not the reason)

Probability of Informed Trading

- ?: use GM-type model to estimate the prob. of informed trading
- On any given trading day, w.p. α there is some information event
  - good or bad
  - observed by informed traders, not by dealer
- Within the day, {informed traders, uninformed sellers, uninformed buyers} arrive as Poisson process with intensity $\{\epsilon_i, \epsilon_s, \epsilon_b\}$
  - Probability of observing n traders of a particular type over the trading day is

$$e^{-\epsilon} \frac{\epsilon^n}{n!}$$

Probability of Informed Trading

- Data for the number of buys and sells allow for estimation of these parameters with maximum likelihood
- ? estimate the 'probability of informed trading' associated with any given order

$$PIN = \frac{\alpha \epsilon_i}{\epsilon_b + \epsilon_s + \alpha \epsilon_i} \tag{5.27}$$

- In NYSE data from 1983 to 1998:
  - median PIN = 19%
  - for 90% of stocks, PIN is between 10% and 30%
  - greater for small-cap stocks, positively correlated with spread and price volatility
- ? show that PIN is higher when markets are more anonymous

Conclusion

Using the insights from previous lectures we can estimate the importance of different components of the spread

- Perhaps surprisingly, order costs are by far the largest cost (but estimated on major stocks)
- Adverse selection is a smaller, yet significant factor
    - is more of an issue in small-cap stocks

Homework

- Read the Economist article on the corporate bond market. Discuss the following questions:
    1. How does corporate bond market liquidity differ from the stock market liquidity? Why?
    2. Why do investors' liquidity expectations matter?
    3. How do investors form their expectations of liquidity?
    4. Can we *measure investors' expectations* of liquidity?
- Fill in the midterm evaluations if you have any comments (link on absalon)
- Start looking at problem set 1

Next week

- Analyze limit order markets: what is the difference to what we have done so far?
- Talk about the role of the 'ticks', the priority rule, the interaction between dealers and LOBs, and a way to interpret limit orders

Derivation of (5.21)

Suppose $q_t = \phi q_{t-1} + \eta_t$.

- **Adverse selection.** Notice that the market valuation now evolves as

$$\mu_t = \mu_{t-1} + \lambda[q_t - \mathbb{E}[q_t|\Omega_{t-1}]] + \epsilon_t.$$

  (Only non-expected part of order flow carries information.)

- **Price.** The price equation then becomes

$$p_t = m_t + \beta q_t + \lambda[q_t - \mathbb{E}[q_t|\Omega_{t-1}]] + \gamma d_t,$$

  where $m_t = \mu_t - \beta z_t$ is the 'mid-price' from the Stoll model (or GM/Kyle models with inventory risk)

Derivation of (5.21)

- Use $q_t = \phi q_{t-1} + \eta_t$ to get $\mathbb{E}[q_t | \Omega_{t-1}] = \phi q_{t-1}$.
- Take first differences of price equation

$$\Delta p_t = \Delta m_t + \lambda[\Delta q_t - \phi \Delta q_{t-1}] + \beta \Delta q_t + \gamma \Delta d_t$$

- Notice that $\Delta z_t = -q_{t-1}$ from market clearing condition. Then

$$\Delta m_t = \Delta \mu_t - \beta \Delta z_t = (\lambda + \beta) q_{t-1} - \lambda \phi q_{t-2} + \epsilon_t.$$

- Substitute in to get

$$\Delta p_t = (\lambda + \beta) q_t - \lambda \phi q_{t-1} + \gamma \Delta d_t + \epsilon_t.$$

Back