Modeling Trade Direction

Dale W.R. Rosenthal

UIC Finance
Liautaud Graduate School of Business

7 March 2009

Motivation

- Financial markets trades result from two or more orders.
- Later arriving order: the *initiator* (aggressor).
- Was the initiator a buy or a sell? aka
  - What was the initiating trade's *direction*? *sign*? *side*?
- Needed for some microstructure research, *e.g.* price impact.
  - Trades impart a small bias (impact) to the price process.
  - Price impact modeled as function of trade size and direction.
- Initiating side (buy/sell) is not available in real-time.
- Fitting impact models is hard, can save $ billions/year.
- Want to guess initiator as accurately as possible.

Current Thinking

- Lee and Ready (1991) first considered delays.
  - Compare trade to midpoint with earlier timestamp.
  - 1988 NYSE and AMEX data: 5 seconds; 1987 data: 2 seconds.
  - Resolve ties with tick test (+ and 0+ ticks: buys)
- The current debate: What method and delay to use?
  - Midpoint test: Vergote (2005) 2s; Henker and Wang (2006) 1s.
  - Bid/ask test, 0s: Ellis, Michaely, and O'Hara (2000); Peterson and Sirri (2003).
  - Tick test: Finucane (2000).

UIC Liautaud

Problems with Previous Studies

Previous work on trade signing has some problems:

Old Data	Pre-electronic, pre-decimalization trades. 1987 (Lee and Ready) to 1999 (Henker and Wang).
Narrow Data	Trades for only a few stocks. 144 (TORQdb) to 401 (Henker and Wang).
Biased Data	Only large-cap stocks (*all* preceding studies).
Time Skew	No simultaneous analyses of NYSE, Nasdaq trades.
Polluted?	Some now-common problems affect many studies.

Why care? "This delay is decreasing to nearly 0 seconds."
Still a problem: delay decreased, but quote volume increased.

Better Quotes and a Modeled Approach

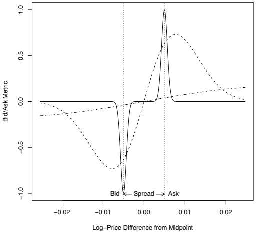
- Picking the correct prevailing quotes may be noisy.
  - Instead, try to get *close* to the prevailing quote.
- Average quotes across time via approximate delay distribution.
- Also use an approach that allows for richer models:
  - Include other information (*e.g.* tick test, bid/ask test);
  - Account for information strengths;
  - Allow for auto-correlated and cross-correlated buys/sells.
  - Acknowledge differences in markets (*e.g.* NYSE vs. Nasdaq).
  - Accommodate effects of market capitalization, liquidity, etc.
- Model can even estimate probability of correct prediction.

UIC Liautaud

Model Notation

$$
\begin{aligned}
b_t, a_t, m_t &= \text{bid, ask, midpoint initiator saw at time } t. \\
p_t &= \text{price of trade at time } t; \\
p_{t-} &= \text{price of trade preceding time } t; \\
p'_{t-} &= \text{differing trade price preceding time } t; \\
B_t &= \text{side of trade at time } t \text{ (1=buy, 0=sell)}; \\
g &= \text{normalized difference function, } \textit{e.g.} \log(p_t) - \log(\hat{m}_t); \\
J &= \text{signed indicator-like function (-1,+1 if } p_t \approx \hat{b}_t, \hat{a}_t\text{; 0 else)}.
\end{aligned}
$$

J needed: estimated quotes may not be decimalized.



J for 1% spread; τ: — 0.1% - - 1% – . 5%

UIC Liautaud

Trade Direction Model

$$P(B_{jt} = \text{Buy}|\mathcal{F}_t; \theta_o, c_k, d_{k\ell}) = \pi_{jt} = \text{logit}(\eta_{jt})$$

$$\eta_{jt} = \underbrace{\beta_0}_{\substack{\text{bias}\\ 0?}} + \underbrace{\beta_{o1} g(p_{jt}, \hat{m}_{jt})}_{\text{midpoint test}} + \underbrace{\beta_{o2} g(p_{jt}, p'_{jt-})}_{\text{tick test}} + \underbrace{\beta_{o3} J(p_t, \hat{b}_t, \hat{a}_t)}_{\text{bid/ask test}} + \underbrace{\phi_o \eta_{jt-}}_{\substack{\text{AR}\\ \text{effect}}} + \underbrace{c_k}_{\substack{\text{overall}\\ \text{effect}}} + \underbrace{d_{k\ell}}_{\substack{\text{within-}\\ \text{sector}\\ \text{effect}}} \quad (1)$$

j indexes stocks; k indexes ten-minute time "bins";
ℓ indexes sectors; o indexes markets.

- Random effects: handle (+) correlations, pseudoreplication.
- Instead of $\phi_o \eta_{jt-}$ AR term, used lagged metrics.

Dataset and Estimation

- Use ArcaTrade dataset from NYSE Archipelago ECN.
  - Includes initiating side for NYSE, Nasdaq, and AMEX stocks[1].
  - Universe: 2,836 different stocks (2004 "Russell 3000").
  - Dec 2004: 1, 2 for estimation; 3–31 for out-of-sample testing.
- In-sample estimation uses almost 2.2 MM observations.
- Out-of-sample testing uses 16.5 MM observations.
- Nonlinear parameters found by conjugate direction (CD).
  - CD uses loop: try parameters, estimate quotes, fit GLMM.
  - Penalized quasi-likelihood used to fit GLMM.

[1] Volume share: 2.3%, 22.5%, 23.3% of NYSE, Nasdaq, AMEX.

Estimated Model

Fixed Effect	AMEX	Nasdaq	NYSE
J width τ	Overall: 2.1×10^{-4} (0.3)		
Delay scale ν	1.66 (0.58)	1.65 (0.65)	0.62 (0.47)
Delay rate λ	0.35 (3.7)	0.33 (0.40)	0.78 (0.35)
Intercept	Overall: 0.06 (0.02)		
Midpoint	—	209 (11)	122 (13)
Tick	—	29.4 (8.4)	-20.5 (8.5)
Bid/Ask	1.20 (0.25)	1.41 (0.02)	2.04 (0.20)
Prev. Bid/Ask	0.33 (0.31)	-0.14 (0.01)	-0.17 (0.05)

Random Effect	Std. Dev.
Time Bin	0.08 (0.01)
Sector $\times$ Time Bin	0.27 (0.03)
Overdispersion Parameter: 1.0086	

UIC Liautaud

Estimation Summary

- Negative prior bid/ask coefficient: agrees with bid-ask bounce.
- Opposite tick coefficient signs: differing short-sale price tests?
- Random effects non-zero, imply buying/selling correlation of:
  - 0.2% across all stocks in 10-minute period.
  - 2% across same-sector stocks in 10-minute period.
- Delay parameter fitting preferred old quotes (30s–120s)
  - Indicates ultra-short-term persistence of quote changes.
- Overdispersion parameter not of practical concern.

UIC Liautaud

Out of Sample: Across Markets

Market	N	Percent of Trades Correctly Classified Modeled	EMO	LR	Tick
AMEX	19,435	69.8%	70.3%	59.2%	52.5%
Nasdaq	15,220,579	74.3%	72.3%	71.8%	66.7%
NYSE	1,264,866	80.7%	79.6%	76.1%	60.7%
Overall	16,504,880	74.7%	72.8%	72.1%	66.2%

EMO = Ellis, Michaely, and O'Hara bid/ask test.
LR = Lee and Ready midpoint test.
Tick = tick test.

- Shocker: LR is the current "gold standard".

UIC Liautaud

Out of Sample: Across Sectors, Spread, Time

- Sectors: Best method across all sectors except one (small).
- Spread: Best method across spread with two exceptions:
  - 0.1% less accurate for 4.4MM trades at ask; and,
  - More abysmal than winner[2] for 30,000 trades at midpoint.
- Dates: Best method for each out-of-sample date.

[2] 45.5% vs. 48.8%.

Performance Attribution: Results

Market	N	Baseline (All Tests)	Convert Tests to Metrics	Add Lag-1 Metrics	Ad-hoc Delay	Full Model
		Change in Percent of Trades Correctly Classified				
AMEX	19,435	67.7%	+2.5%	+0.4%	-0.8%	+0.0%
Nasdaq	15,220,579	70.3%	+3.0%	-0.1%	+0.9%	+0.2%
NYSE	1,264,866	79.8%	+1.1%	-0.6%	+0.7%	-0.3%
Overall	16,504,880	71.1%	+2.7%	-0.1%	+0.9%	-0.1%

- To attribute performance, I fit a series of nested models.
- Information strength (± 1 tests to metrics) gains 1%–3%.
- Adding lagged bid/ask metric gains 0.4% for AMEX trades.
- Basic delay model gains 0.8% for NYSE, Nasdaq trades.

UIC Liautaud

Contributions and Further Work

- Beat next-best method by 1–2%[3] across almost all groupings.
- Introduced delay theory to estimation of prevailing quotes.
- Opened doors to richer trade signing models:
  - Use multiple sources of information.
  - Consider strength of information.
  - Correct for microstructure peculiarities.
  - Allow for autocorrelations and cross-correlations.
  - Interaction between volume/volatility/spread and metrics?
- Shown short- and ultra-short-term buying/selling persistence.
- Developed Edgeworth expansions for average delays.
- Conduct experiments to infer BLUPs and make money?

[3] *cf.* Most published results beat EMO or LR by 0.5% in one group.