# Modeling Trade Direction

Dale W.R. Rosenthal (daler@uic.edu)
*Department of Finance, University of Illinois at Chicago*


## Summary

*I propose a modeling approach to classify trades by whether buyer or seller demanded liquidity.*

*The model (i) uses estimated quotes in midpoint, tick, and bid/ask tests; (ii) considers test strengths; (iii) can account for microstructure effects; and, (iv) allows for serial and cross-correlations in trade direction.*

*A recent sample of trades for 2,836 US stocks is used to compare the model to other classification methods. Out of sample, modeled classifications are 1–2% more accurate overall than current methods. This improvement is consistent across dates, sectors, and the spread.*

*For Nasdaq and NYSE stocks, 0.9% and 0.7% of the improvement comes from estimating quotes; 1% and 1.3% of the improvement comes from using the strengths of the various tests. For AMEX stocks, a lagged version of the bid/ask test accounts for 0.4% of improvement.*

*I also find indications of short- and ultra-short-term alpha.*


## Introduction

### Trade Classification

Lee and Ready (1991) analyzed NYSE trades to infer the liquidity taker ("initiator") by comparing trade prices to quote midpoints prevailing at trade time (the "LR method"). They also noted trade reports were delayed and suggested using 5-second-lagged quotes. Vergote (2005) and Henker and Wang (2006) considered other lags for quotes.

Ellis, Michaely, and O'Hara's (2000) analysis of Nasdaq trades compared trade prices to unlagged bids and asks (the "EMO method"). Peterson and Sirri (2003) recommended the EMO method for classifying NYSE stock trades.

Both LR and EMO methods resolve indeterminacy by comparing trade prices to preceding trade prices ("tick test"). Finucane (2000) suggested classifying trades with only a tick test.

Caudill, Marshall, and Garner (2004) tried and failed to model trade direction.

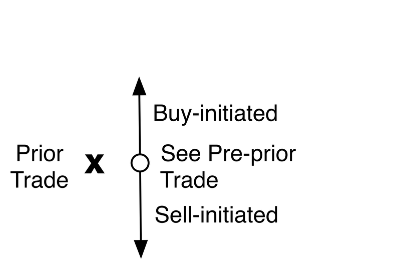
### Delays and Small-scale Asymptotics

Delay theory dates from Erlang (1903). Lee and Ready (1991) first considered trade reporting delays. I derive non-standard (gamma-based) Edgeworth expansions to approximate delay distributions à la McCullagh (1987).

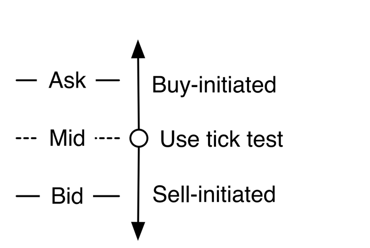
## Why Trade Classification Matters

Trade classification is used to estimate net order flow as well as models of market impact, the probability of informed trading, and effective spreads. These models can save billions of dollars in trading costs — and may be used to infer alpha.

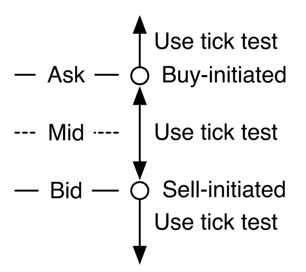
## Current Trade Classification Methods




Tick Test




LR Method
(a midpoint test)

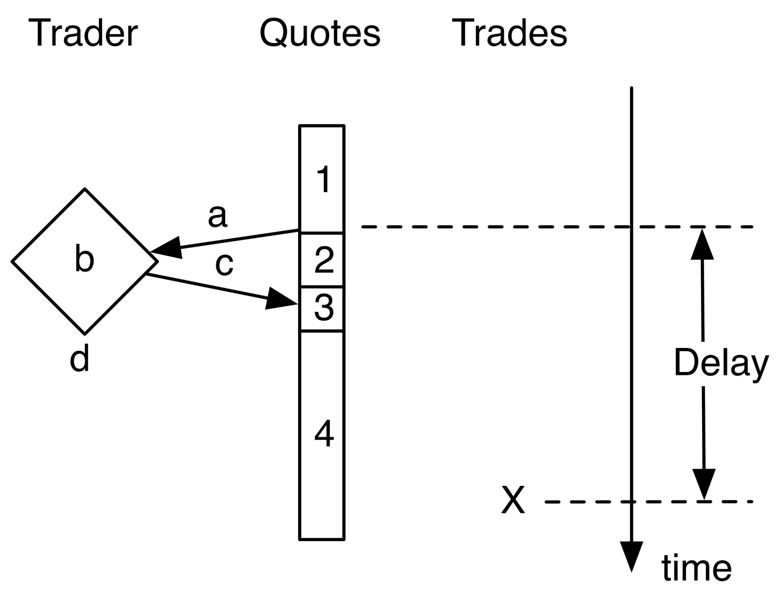



EMO Method
(a bid/ask test)

| | Tick Test | LR Method | EMO Method |
|---|---|---|---|
| Patron: | Finucane (2000) | Lee and Ready (1991) | Ellis, Michaely, and O'Hara (2000) |
| Buy: | Trade Price > Prior Trade | Trade Price > Midpoint | Trade Price = Ask |
| Sell: | Trade Price < Prior Trade | Trade Price < Midpoint | Trade Price = Bid |
| Else: | Refer to pre-Prior Trade | Use Tick Test | Use Tick Test |

The Stevanovich Center for Financial Mathematics at the University of Chicago provided computing resources; financial support provided by National Science Foundation grants DMS 06-04758 and SES 06-31605.

## Example of Delay Between Quote and Trade Timestamps




1) Message a: Quote #1 received by trader.
2) Decision b: Trader decides to trade; assembles order.
3) Message c: Order sent to market center.
4) Trade occurs; quote updated quickly (quote #3 → #4).
5) Private message d: Trader notified of fill.
6) Trade ("X") printed (and timestamped) later.

## Problems with Previous Studies

**Old Data** Used US trades from pre-electronic, pre-decimal era: 1987 (LR), 1990 (multiple), 1997 (EMO, Peterson and Sirri), 1999 (Henker and Wang).

**Narrow Data** Used few stocks: 144 (TORQdb), 150 (LR), 313 (EMO), 401 (Henker and Wang).

**Biased Samples** Used only large-cap stocks (*all* preceding studies).

**Ad-hoc Delay** Guessed prevailing quotes via deterministic lag: 5s (LR), 2s (Vergote), 1s (Henker and Wang), 0s (EMO, Peterson and Sirri).

**Time Skewed** Did not analyze contemporaneous data for Nasdaq and NYSE stocks.

**Ignored Information Strength** Exceedance of test thresholds not considered. For example: trades \$0.01 and \$1.00 above the midpoint were both seen as definitely buy-initiated.

**Possibly Polluted** Certain increasingly common negotiated trades (VWAP, PRP) may have biased previous (and current) studies.

**Solution** Use Dec 2004 dataset of 2.2 million non-negotiated trades in 2,836 small- to large-cap stocks across US markets (AMEX, Nasdaq, NYSE). Convert price tests to distance-like *metrics*; use a delay distribution to estimate prevailing quotes; let price metrics "vote" on classification.

## Metrics

| | |
|---|---|
| $g(p_t, p'_{t-}) = \log(p_t) - \log(p'_{t-})$ | compares trade price to prior differing trade price |
| $g(p_t, m_t) = \log(p_t) - \log(m_t)$ | compares trade price to midpoint. |
| $J(p_t, b_t, a_t) = e^{-(\frac{p_t - a_t}{\tau})^2} - e^{-(\frac{p_t - b_t}{\tau})^2}$ | measures trade price proximity to bid or ask. |

## Models

**Delay** A gamma-based Edgeworth expansion was tried; higher order terms were not significant. Thus random delay $Y \sim \mathrm{Gamma}(\nu, \lambda)$. The estimated ask $\hat{a}_t$ for a trade reported at time $t$ is then: $\hat{a}_t = \int_0^\infty \gamma_{\hat{\nu},\hat{\lambda}}(y) a_{t-y} dy$.

**Classification** A generalized linear mixed model of the likelihood a trade was buyer-initiated:

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

$$\eta_{jt} = \underbrace{\beta_0}_{\substack{\text{bias}\\=0?}} + \underbrace{\beta_{o1} g(p_{jt}, \hat{m}_{jt})}_{\text{midpoint metric}} + \underbrace{\beta_{o2} g(p_{jt}, p'_{jt-})}_{\text{tick metric}} + \underbrace{\beta_{o3} J(p_{jt}, \hat{b}_{jt}, \hat{a}_{jt})}_{\text{bid/ask metric}} +$$
$$\underbrace{\beta_{o4} g(p_{jt-}, \hat{m}_{jt-})}_{\text{lag-1 midpoint metric}} + \underbrace{\beta_{o5} g(p_{jt-}, p'_{jt--})}_{\text{lag-1 tick metric}} + \underbrace{\beta_{o6} J(p_{jt-}, \hat{b}_{jt-}, \hat{a}_{jt-})}_{\text{lag-1 bid/ask metric}} + \underbrace{c_k}_{\substack{\text{overall}\\\text{effect}}} + \underbrace{d_{k\ell}}_{\substack{\text{within-}\\\text{sector}\\\text{effect}}} .$$

with indices $j$ (stocks), $k$ (time bins), $\ell$ (sectors), and $o$ (markets) and random effects $c_k \overset{\text{iid}}{\sim} (0, \sigma_c^2)$, $d_{k\ell} \overset{\text{iid}}{\sim} (0, \sigma_d^2)$.

## Fixed and Random Effect Estimates

| Fixed Effect | AMEX | Nasdaq | NYSE |
|---|---|---|---|
| $\tau$ | Overall: $2.1 \times 10^{-4}$ (0.3) | | |
| $\nu$ | 1.66 (0.58) | 1.65 (0.65) | 0.62 (0.47) |
| $\lambda$ | 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 × Time Bin | 0.27 (0.03) |

- Trading at the bid/ask appears to be highly informative.
- Nasdaq and NYSE tick terms differ in sign, reflecting different short sale tests.
- Negative coefficients for previous bid/ask term reflects bid-ask bounce.

**Indications of Alpha** A few details indicate short- to ultra-short-term alpha:

- Time random effect ⇒ 0.2% buy/sell correlation over all stocks in 10-minute period.
- Sector×time random effect ⇒ 2% buy/sell correlation over sector stocks in 10-minute period.
- Unconstrained delay parameters ⇒ ultra-short-term buy/sell persistence (over 30s–120s).

## Out of Sample Testing: Trades Correctly Classified

**By Market**

| Market | N | Modeled | EMO | LR | Tick |
|---|---|---|---|---|---|
| | | Percent of Trades Correctly Classified | | | |
| 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% |

**Across Estimated Bid-Ask Spread** Model outperforms other methods across the spread with two exceptions: (i) the model is 0.1% less accurate for 4.4MM trades at the ask, and (ii) the model is more abysmal than the winner (45.5% vs. 48.8%) for 30,000 trades at the midpoint.

Contrary to previous findings, classification outside the spread is more accurate — perhaps due to excluding negotiated trades.

### Attributing Performance to Model Features

| Market | N | Baseline (All Tests) | Converting Tests to Metrics | Adding Lagged Metrics | Including 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% |

## Conclusion

Modeled trade classifications outperform all other classification methods across major markets and across almost all parts of the spread.

A modeling approach allows accounting for interactions, microstructure effects, and correlations; also, this approach changes the debate from choosing methods to incorporating all methods.

Finally, I see nonparametric evidence of return predictability over short (10-minutes) and ultra-short (30s–120s) time periods. Experiments to infer BLUPs may allow this alpha to be used.

## References


1. Caudill, S. B.; B. B. Marshall and J. Garner. "Improved Trade Classification Rules: Estimates Using a Logit Model Based on Misclassified Data", *Atlantic Economic Journal*, 32:3(2004), 256.
2. Ellis, K.; R. Michaely and M. O'Hara. "The Accuracy of Trade Classification Rules: Evidence from Nasdaq", *Journal of Financial and Quantitative Analysis* 35:4(2000), 529–551.
3. Finucane, T. J. "A Direct Test of Methods for Inferring Trade Direction from Intra-Day Data", *Journal of Financial and Quantitative Analysis* 35:4(2000), 553–576.
4. Henker, T. and Wang, J. "On the Importance of Timing Specifications in Market Microstructure Research", *Journal of Financial Markets* 9(2006), 162–179.
5. Lee, C. M. C. and Ready, M. J. "Inferring Trade Direction From Intraday Data", *Journal of Finance* 46:2(1991), 733–746.
6. McCullagh, P. *Tensor Methods in Statistics*, 1987. Chapman and Hall: London.
7. Peterson, M. and Sirri, E. "Evaluation of the Biases in Execution Cost Estimation Using Trade and Quote Data", *Journal of Financial Markets* 6(2003), 259–280.
8. Vergote, O. "How to Match Trades and Quotes for NYSE Stocks?" KU Working Paper. Katholieke Universiteit Leuven (2005).