Lack of Anonymity and the Inference from Order Flow

Juhani Linnainmaa and Gideon Saar
Central Bank Workshop on the Microstructure of Financial Markets, October 7, 2010
Motivation

- The objective of the paper is to contribute to the literature that investigates whether market structure rules that govern trading information are important.
  - Our focus: Information about the identity of investors.

- The concept of “Anonymity” in financial markets pertains to the information market observers have about the identity of investors who submit orders.
  - Intermediate level: ELOB with broker ID; floor markets.
  - Non-anonymous: Upstairs markets.

- Degree of Anonymity is important: Event studies find that liquidity can change when broker ID is introduced or eliminated.
Motivation – cont.

◆ Why is broker ID important?
  – Informed investors could use mixed strategies (i.e., trade through multiple brokers) in a manner that makes this signal uninformative.
  – In the absence of frictions in the economic environment, intermediate level of anonymity is irrelevant.

◆ Our goal: To investigate the information content of signals about the identity of investors and whether they affect price formation in a market with an intermediate level of anonymity.
Data and Sample

- Two sources of data:
  - Finnish Central Securities Depository registry (complete trading records of all Finnish investors).
  - Helsinki Stock Exchange supervisory files (information on every order that is entered into the exchange’s system).

- The trading mechanism of the Helsinki Stock Exchange is an electronic limit order book with standard price-time priority rules.

- Key feature: broker IDs are visible.


- Sample: 87 firms (all listed firms with average daily number of trades exceeding five).
  - All 41 brokerage houses that trade these stocks.
Investor Types

- Investors are grouped into three types:
  - Domestic households
  - Foreign investors
  - Domestic institutional investors

- Maintained assumption: Domestic institutions are more informed on average than domestic households (i.e., they possess more pricing-relevant skill or information about the stocks).
  - Results on foreigners are documented as stylized facts rather than used to test hypotheses about anonymity.
Broker Assignment Rule

- A broker belongs to a Broker Group (BG) that is associated with a certain investor type if more than 50% of the trades that the broker executes involve this investor type.
  - BG1 associated with households, BG2 associated with foreigners, BG3 associated with domestic institutions.

- We test the pricing implications of inference from broker identities, not how prices respond to the trading of particular investor types.

<table>
<thead>
<tr>
<th>Broker Group</th>
<th>Investor Category</th>
<th>Households</th>
<th>Foreigners</th>
<th>Domestic Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG1</td>
<td></td>
<td>72.3%</td>
<td>6.9%</td>
<td>20.8%</td>
</tr>
<tr>
<td>BG2</td>
<td></td>
<td>3.7%</td>
<td>79.1%</td>
<td>17.2%</td>
</tr>
<tr>
<td>BG3</td>
<td></td>
<td>11.5%</td>
<td>16.9%</td>
<td>71.6%</td>
</tr>
</tbody>
</table>
Absence of frictions associated with broker selection, informed investors would want to use mixed strategies across the different brokers to “hide” their order flow.

 Majority of institutional trading comes from multi-broker users!
- Consistent with informed institutions that are aware of the information content of their order flow and attempt to “hide.”
Inference about Investor Types

- Is there still an information content to broker ID?
- Probit regressions (pooled; stock fixed-effects).
  - Dependent variable in the Households regression:
    » 1=if a household submits a marketable order, 0=otherwise.
  - Controls: recent activity (volume, signed return, volatility, duration), prevailing state of the limit order book (BBO depth; spread) and trade size (trade size; size-relative-to-depth).

<table>
<thead>
<tr>
<th></th>
<th>Households</th>
<th>Foreigners</th>
<th>Domestic Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG1 Coef.</td>
<td>0.2123**</td>
<td>-1.3568**</td>
<td>-0.4896**</td>
</tr>
<tr>
<td>t-stat.</td>
<td>(21.29)</td>
<td>(-120.07)</td>
<td>(-54.41)</td>
</tr>
<tr>
<td>BG2 Coef.</td>
<td>-2.1319**</td>
<td>0.8487**</td>
<td>-0.7838**</td>
</tr>
<tr>
<td>t-stat.</td>
<td>(-204.92)</td>
<td>(74.88)</td>
<td>(-85.77)</td>
</tr>
<tr>
<td>BG3 Coef.</td>
<td>-1.5413**</td>
<td>-0.9872**</td>
<td>0.8795**</td>
</tr>
<tr>
<td>t-stat.</td>
<td>(-143.85)</td>
<td>(-85.31)</td>
<td>(93.79)</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Controls: No Yes No Yes No Yes
Inference about Investor Types–cont.

- How much information is there in broker ID?
  - When broker group dummies are added to the regressions, the pseudo-$R^2$ increases:
    - Households regression: from 25.06% to 43.95%
    - Foreigners regression: from 5.66% to 30.45%
    - Institutions regressions: from 2.58% to 15.77%

- Result: Broker ID can be used by market participants to significantly increase their ability to infer who (in terms of investor types) is behind initiated trades.
  - It provides an informative signal even though the majority of trading comes from multi-broker users.
Permanent Price Impact

- Is broker ID information meaningful enough to affect price formation in the market?
- The Permanent Price Impact of a trade measures price adjustment from an instant before the arrival of the marketable order to a time where we assume prices have finished their adjustment to the information content of the order.
- We use the signed log change in the midquote from an instant before the trade to five minutes after the trade.
  - Results robust to using other definitions of permanent price impact.
  - We subtract the average price impact in the same stock for marketable orders in the same direction.
### Information Asymmetry Groups

- **Result 1:** Prices adjust more to marketable orders from brokers associated with better-informed investors.
- **Result 2:** Degree of information asymmetry affects price adjustment for orders coming through brokers associated with informed investors, but not for orders from brokers associated with uninformed investors.

<table>
<thead>
<tr>
<th>Broker Group</th>
<th>Bid-Ask Spread Quartiles</th>
<th>t-test((Q_4-Q_1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG1</td>
<td>0.069</td>
<td>0.052</td>
</tr>
<tr>
<td>BG2</td>
<td>0.017</td>
<td>0.117</td>
</tr>
<tr>
<td>BG3</td>
<td>0.031</td>
<td>0.134</td>
</tr>
</tbody>
</table>

- \(t\)-test\((BG1-BG2)\): -72.8, -33.0, -25.1, -10.8
- \(t\)-test\((BG1-BG3)\): -59.7, -26.7, -24.6, -14.6
- \(t\)-test\((BG2-BG3)\): -10.1, -2.1, -1.3, -2.7

- \(t\)-test \(BG1_{(Q_4-Q_1)}-BG2_{(Q_4-Q_1)} = -7.3\)
- \(t\)-test \(BG1_{(Q_4-Q_1)}-BG3_{(Q_4-Q_1)} = -10.4\)
- \(t\)-test \(BG2_{(Q_4-Q_1)}-BG3_{(Q_4-Q_1)} = -2.2\)
Permanent Price Impact Regressions

- This is where we try to hold everything else equal and see whether otherwise-identical trades have larger permanent price impacts when they come through a broker that is associated with better informed investors.

- Pooled regressions with stock fixed-effects.
  - Orders coming from BG1 brokers serve as the (unreported) stock-specific intercept.

\[
ppi_{i,t} = a_i \text{StockDummies}_{i,t} + b_1 \text{BG2}_{i,t} + b_2 \text{BG3}_{i,t} + b_3 \text{Volume}_{i,t} + b_4 \text{SgnReturn}_{i,t} + \\
+ b_5 \text{Volatility}_{i,t} + b_6 \text{TradeSize}_{i,t} + b_7 \text{TradeSize}_{i,t}^2 + b_8 \text{SameSideDepth}_{i,t} + \\
b_9 \text{OtherSideDepth}_{i,t} + b_{10} \text{FirstTrade}_{i,t} + b_{11} \text{Duration}_{i,t} + b_{12} \text{Duration}_{i,t}^2 + \\
+ b_{13} \text{Spread}_{i,t} + b_{14} \text{Spread}_{i,t}^2 + b_{15} \text{MarketTrade}_{i,t} + b_{16} \text{LargeTrade}_{i,t} + \epsilon_{i,t}
\]
## PPI Regressions: Results

- Similar results to those in the information asymmetry groups!

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Market Capitalization</th>
<th>Bid-Ask Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Q1(small)</td>
<td>Q2</td>
</tr>
<tr>
<td>BG2</td>
<td>Coef.</td>
<td>0.040**</td>
<td>0.106**</td>
</tr>
<tr>
<td></td>
<td>t-stat.</td>
<td>(23.08)</td>
<td>(5.60)</td>
</tr>
<tr>
<td>BG3</td>
<td>Coef.</td>
<td>0.067**</td>
<td>0.126**</td>
</tr>
</tbody>
</table>
Interactions

- There are several interesting interactions that demonstrate the sophisticated manner in which market participants make a joint inference from broker ID and other order attributes.

- Example: Duration
  - Unconditional effect on PPI < 0 (Easley and O’Hara (1992), Dufour and Engle (2000)).
  - BG1 * Duration > 0 (more “noise” trading)
  - BG3 * Duration < 0 (reinforces unconditional effect)
Conclusions

We show that despite the utilization of multiple brokers by some investors, broker ID can still be used as a powerful signal to help classify orders into “investor types.”

Implication: There must be frictions in the economic environment that prevent investors from sending orders through multiple brokers.

- Possible frictions: Heterogeneity in brokerage fees; quantity discounts; bundling of services offered by various brokers (e.g., research; Order Management Systems); “preferred customer” effort level.

Goldstein, et al. (2009): Bundling in brokerage industry prevents order flow from going to the broker with the lowest commission.

We: Significant frictions prevent informed investors from migrating to the brokers that would afford them the lowest price impact of trading.
We show that signals about the identity of investors (i.e., broker ID) indeed make a difference with respect to market prices.

Our findings explain why event studies of changes in anonymity show a significant impact on market liquidity.
  - Typical finding is that the removal of broker ID helps liquidity.
  - Our study points to a less favorable interpretation of this result: improved liquidity arises because informed investors are able to hide better and therefore comes at the expense of informational efficiency.

Current developments in trading technology (e.g., algorithmic trading) should increase the utilization of signals such as broker identity.