Measuring the Amount of Asymmetric Information in the Foreign Exchange Market

Esen Onur[1] and Ufuk Devrim Demirel[2]

September 2009

VERY PRELIMINARY & INCOMPLETE

PLEASE DO NOT CITE WITHOUT AUTHORS' PERMISSION

Abstract

Does every investor make use of the same amount of information while they trade in the foreign exchange market? We do not believe they do. The aim of this paper is to measure the amount of asymmetric information in a specific foreign exchange market using a unique data set from the Bank of Israel. We make use of a microstructure approach built on a structural sequential trade model referred to as the *PIN Model* (Probability of information-based trading) in finance. This model measures the probability that a customer is more or less-well informed given a buy or a sell trade is initiated by that customer in the market. Our results show that in the Israeli foreign exchange market, the probability of a trade being "informed" is much higher for a certain group of investors, *foreign financials*. For the other group of investors, *other customers*, this probability is significantly lower.

Keywords: Foreign Exchange, Asymmetric Information, PIN

[1] Corresponding author: Esen Onur, Department of Economics, California State University, Sacramento.
[2] Ufuk Devrim Demirel, Department of Economics, University of Colorado – Boulder.

Introduction:

One of the important questions in the foreign exchange (FX) market is how much information do investors in the market really possess? The answer definitely depends on which market one is interested in, as well as the time interval considered. On top of that, it is impossible to even consider a complete FX market due to the decentralized aspect of it. There are many dealers located all around that world who engage in purchase and sale of currency pairs. Thus, without a market maker or a specialist, it is not possible to collect information about every transaction that takes place between a given currency pair.

With all of these limitations in mind, in that sense, we have access to a daily data set from the Israeli FX market from 1999 to June of 2006. The data set is collected by the Central Bank of Israel (BOI) and these are daily numbers reported to BOI by commercial banks operating in Israel. The currency pair we have access to is the New Israeli Sheqel / US Dollar. Even though we are not covering the whole market with this data set, we think a big proportion of this currency pair is traded within the borders of Israel and are captured in our data set.

Empirical literature tells us that every investor does not possess the same amount of quality of information. Many studies have found empirical evidence to suggest that order flow[3] from different parties such as big or small banks, foreign or domestic banks, leveraged or non-leveraged companies, all have different impact on the price (equilibrium exchange rate)[4]. These findings are encouraging but not conclusive about how much private information these parties possess.

On the other hand, equity markets have developed a statistical way to measure the probability of an order flow from a certain kind of investor having more information than other investors, the PIN technique. We borrow this methodology to apply it to our foreign exchange market data and measure the probability that different kinds of agents in our data set might have different probabilities of private information.

Our results show that in the Israeli foreign exchange market, the probability of a trade being "informed" is much higher for a certain group of investors, *foreign financials*. For the other group of investors, *other customers*, this probability is significantly lower. This is true for all the years in our data set but the magnitude of difference in PIN measure between these two groups change from one year to another. We find that it is highest in the year 2004 but also lowest in the following year and a half.

Reports from bank of Israel suggest that the FX market was changing as the customer group they call "foreign financials" were coming into the market between 1999-2002. After 2002, they claim that market has matured. Using our data and this PIN method, we test this claim. If there is really more private information in the market post-maturing (after 2002)[5], then we can claim that this "foreign

[3] Order flow is (net) signed trade. If it is buyer initiated, it is positive and if seller initiated it is negative.
[4] A few of these studies are Onur (2008), Marsh and O'Rourke (2006), Dagfinn (???)
[5] This part may call for testing for a structural break instead of just guessing the year 2002.

financials" customer group does have more private information. This is another way of identifying the group of investors with more information, rather than assuming that they do.

Model:

Setup:

The model used in this paper is a special case of Glosten and Milgrom (1985)[6]. It is also heavily influenced by Easley, Hvidjkaer and O'Hara (2002)[7].

The essential feature of these kinds of models is that trade reveals something about an investor's information. Purchase of currency (a "buy" order) will generally come from an informed investor who has positive information but it will not come from an informed investor with negative information. This kind of rational explains why "buy" or "sell" orders from investors actually carry important information.

Microstructure treats trading as a game between market maker and traders that is repeated over trading days. First, nature chooses whether there is new information at the beginning of the trading day and it occurs with probability α. The new information is signal with regards to the actual return of this security (V_i). Return of this security is either high and worth $\bar{V}_i$, or low and worth $\underline{V_i}$. High return occurs with probability (1-δ) and bad news occurs with probability δ.

Trading for day i starts with traders arriving according to Poisson process throughout the day. The market maker sets the prices to buy (B) or sell (S) and then executes orders as they arrive. Traders, on the other hand, can be informed or uninformed. Let us assume for now that informed traders know the value of the return on this security. The proportion of informed traders in the population is μ. Also, the dealer does not know whether the trader is informed or not.

The sequential trade model is summarized in the following diagram:

---

[6] Glosten, L. R., and Milgrom, R. P., 1985, "Bid, ask, and transaction prices in a specialist market with heterogeneously informed trader," *Journal of Financial Economics* 14, 71-100.
[7] Easley, D., Hvidjkaer, S., and O'Hara, M., 2002, "Is Information Risk a Determinant of Asset Returns?" The Journal of Finance 57, No. 5, 2185-2221.

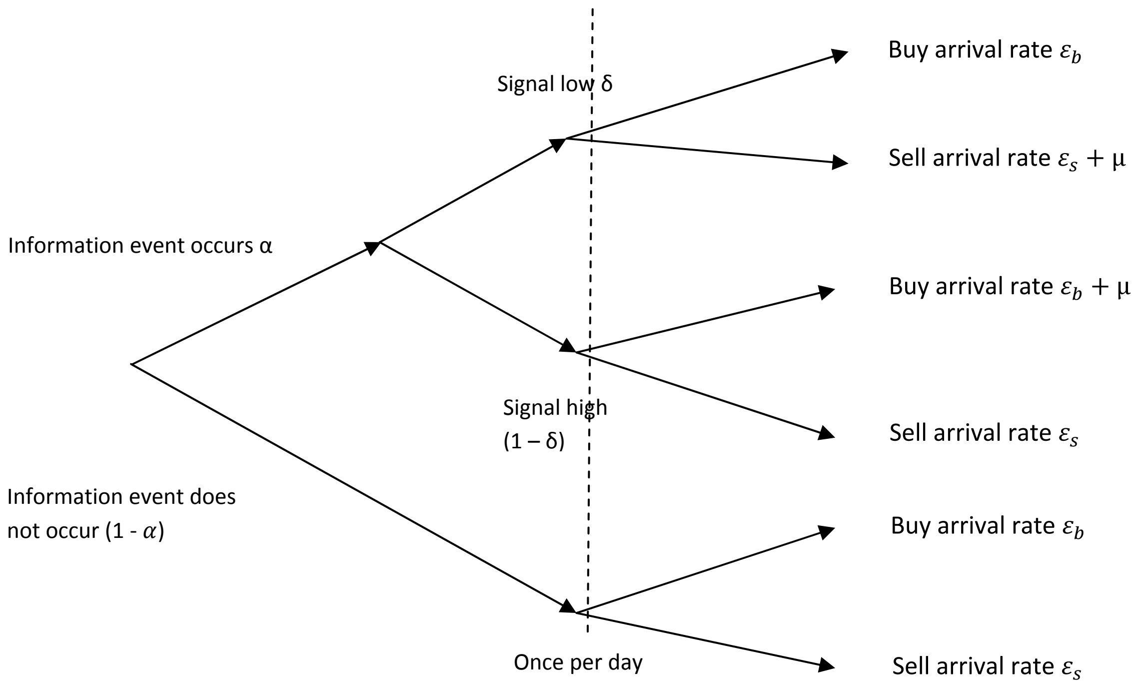


Figure 1 Tree diagram of the trading process

The tree given in figure 1 describes the trading process. At the first node of the tree, nature selects whether an information event occurs. If an event occurs, nature then determines if it is good news (high signal) or bad news (low signal). Good news suggests a high return for the investment and a bad news suggests a low return (loss). Conditional on new information occurring, good news happens with probability (1-δ) and bad news with probability δ. The three nodes (no event, good news and bad news) before the dotted line in figure 1 occur only once per day. Then, given the node selected for the day, traders arrive according to the relevant Poisson process. Market maker, on the other hand, sets buy and sell prices every day and execute orders as they arrive. That is, on good event days, the arrival rates are ε_b + μ for buy orders and ε_s for sell orders. On bad event days, the arrival rates are ε_b for buys and ε_s + μ for sells. Finally, on no-event days, only uninformed traders arrive and the arrival rate of buys and sells are ε_b and ε_s respectively.

The market maker is assumed to be a Bayesian who uses the information in the arrival of trades and their intensity to update his expectations about whether it is good news, a bad news or a no news day. Arrival of daily news is assumed to be independent. This means market maker's analyses are independent from one day to another. At any time *t* before trading starts, market maker's prior beliefs for probability of no news, good news and bad news are the as follows respectively:

$$P_n(t) = 1 - \alpha, \;\; P_g(t) = \alpha(1 - \delta), \;\; P_b(t) = \alpha\delta \qquad (1)$$

Let us use S_t and B_t to denote sell and buy orders at time *t*. The market maker uses the information apparent in the buy or sell orders that arrives to update his beliefs about the market. Let $P(t|S_t)$ to denote the market maker's updated belief conditional on a sell order arriving at time *t*. $P_n(t|S_t)$ is market maker's belief about no news conditional on a sell order arriving at time *t*. Similarly, $P_b(t|S_t)$ is market maker's updated probability of bad news conditional on the fact that a sell order arrives. Finally, $P_g(t|S_t)$ is his updated probability of good news conditional on the fact that a sell order arrives. Probabilities conditional on buy order arriving are represented in a similar fashion.

Market Maker's Updated Beliefs:

Using Bayes's rule, the market maker's posterior probability on no news at time *t*, if a sell order arrives at time *t*, is

$$P_n(t|S_t) = \frac{P_n(S_t|t)P_n(t)}{P(S_t)}$$

$$= \frac{P_n(S_t|t)P_n(t)}{P_n(S_t|t)P_n(t) + P_g(S_t|t)P_g(t) + P_b(S_t|t)P_b(t)}$$

$$= \frac{\varepsilon_s P_n(t)}{\varepsilon_s + \mu P_b(t)} \qquad (2)$$

Similarly, the posterior probability on bad news is

$$P_b(t|S_t) = \frac{(\varepsilon_s+\mu)P_b(t)}{\varepsilon_s + \mu P_b(t)} \qquad (3)$$

and the posterior probability on good news is

$$P_g(t|S_t) = \frac{(\varepsilon_s)P_g(t)}{\varepsilon_s + \mu P_b(t)} \qquad (4)$$

At time *t*, the zero expected profit bid price, *b(t)*, is the market maker's expected value of the asset. Thus, the bid at time *t* is

$$b(t) = \frac{\varepsilon_s P_n(t)V^* + (\varepsilon_s+\mu)P_b(t)\underline{V_i} + (\varepsilon_s)P_g(t)\bar{V}_i}{\varepsilon_s + \mu P_b(t)} \qquad (5)$$

Where V^* denotes the value of the security if no news happens. Similar calculations show that the ask at time *t* is

$$a(t) = \frac{\varepsilon_b P_n(t)V^* + \varepsilon_b P_b(t)\underline{V_i} + (\varepsilon_b + \mu)P_g(t)\bar{V}_i}{\varepsilon_b + \mu P_g(t)} \qquad (6)$$

To put things in perspective, it is useful to relate these bid and ask prices to the expected value of the asset. Expected value of the asset is:

$$E[V_i] = P_n(t)V^*{}_i + P_b(t)\underline{V_i} + P_g(t)\bar{V}_i \qquad (7)$$

Substituting equation (7) into equation (5) and (6) gives us the modified bid and ask prices,

$$b(t) = E[V_i] - \frac{\mu P_b(t)}{\varepsilon_s + \mu P_b(t)}\left(E[V_i] - \underline{V_i}\right) \qquad (8)$$

and

$$a(t) = E[V_i] + \frac{\mu P_g(t)}{\varepsilon_b + \mu P_g(t)}(\bar{V}_i - E[V_i]) \qquad (9)$$

Writing these equations with respect to the expected value of the asset helps us analyze the role played by arrival of informed and uninformed investors in affecting trading prices. If there are no informed traders, $\mu = 0$, then trade carries no information and both bid and ask prices are equal to the expected value of the asset. On the other hand, if buy and sell arrival rates that are not contingent on information were non-existent, $\varepsilon_b = \varepsilon_s = 0$, then $b(t) = \underline{V_i}$ and $a(t) = \bar{V}_i$ for all t. At these prices no informed trades will trade and market shuts down.

Generally, both informed and uninformed traders will be in the market and the bid (ask) will be below (above) $E[V_i]$. Now, let us denote this spread as the difference between ask and bid prices,

$$\Sigma(t) = a(t) - b(t),$$

such that

$$\Sigma(t) = \frac{\mu P_g(t)}{\varepsilon_b + \mu P_g(t)}(\bar{V}_i - E[V_i]) + \frac{\mu P_b(t)}{\varepsilon_s + \mu P_b(t)}\left(E[V_i] - \underline{V_i}\right) \qquad (10)$$

This spread can be analyzed in two parts. The first part is the expected loss to an informed investor times the probability that the buy in the market is information-based. The second part is the expected loss from selling to an informed investor times during bad times the probability that the sale is an information-based one.

How to define PIN?

Let us define PIN as average of these probabilities:

$$PI(t) = \frac{1}{2}\left(\frac{\mu P_g(t)}{\varepsilon_b + \mu P_g(t)} + \frac{\mu P_b(t)}{\varepsilon_s + \mu P_b(t)}\right) \qquad (11)$$

If it was average of probabilities: *The reasoning being that the market maker observes the type of trade, a buy or a sell. Given that it is a buy or a sell, the probability that it is informed is given by the two terms in equation (11). Since there might be many buy and sell orders in any given day, we take the average of these probabilities as the PIN measure.* ***A better way might be to weigh them with the number of buy or sell orders associated with the probability.***

This probability measures the proportion of the arrival ratio of information based trades to the arrival ratio of total trades. For example if there are no informed traders in the market, μ should be equal to zero. That means probability of informed trading will also be zero, PI(*t*)=0. Alternatively, if all trades are based on information, then $\varepsilon_b = \varepsilon_s = 0$. That means PI(*t*)=1.

Likelihood Function:

We need to estimate the parameter vector $\theta = (\alpha,\ \delta,\ \varepsilon_s,\ \varepsilon_{b,}\ \mu)$ to calculate the probabilities in equation (11). Note that these parameters are not really observed individually in the data. To give an example, parameters α and δ determine the probabilities of three information events; no news, good news and bad news, none of which are observable to us. The remaining parameters refer to the arrival rates of informed and uninformed traders. We observe the total magnitude of buy or sell orders but we do not observe which traders are informed or uninformed in the market. To estimate all these parameters, we use a maximum likelihood estimation based on the number of buy and sell orders observed.

In the model, we realize that the data signal the information setup to us through more buy orders expected on days with good events, and more sell orders expected on days with bad events. Similarly, on no-event days, there are no informed traders in the market so fewer orders should arrive. So, the fact that we observe whether we have a buy or a sell order allows us estimate the parameter vector by making use of the likelihood of these orders being observed.

Let us construct the likelihood of all these orders under three different information scenarios. On a bad-event day, sell orders will arrive at the rate of $\mu + \varepsilon_s$. On the other hand, buy orders will arrive at the rate of ε_b suggesting that only uninformed investors would be buying when there is a bad information event. We assumed the distribution of these statistics to be Poisson, so the likelihood of observing any sequence of orders that contain *B* buys and *S* sells on a bad event day is

$$e^{-\varepsilon_b} \frac{{\varepsilon_b}^B}{B!} e^{-(\mu+\varepsilon_s)} \frac{(\mu + \varepsilon_s)^S}{S!} \qquad (12)$$

Similarly, on a good event day, this likelihood is

$$e^{-(\mu+\varepsilon_b)} \frac{(\mu + \varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{{\varepsilon_s}^S}{S!} \qquad (13)$$

Finally, on a no-event day, this likelihood becomes

$$e^{-\varepsilon_b}\frac{{\varepsilon_b}^B}{B!}e^{-\varepsilon_s}\frac{{\varepsilon_s}^S}{S!} \tag{14}$$

The likelihood of observing *B* buys and *S* sells on a day of unknown type is the weighted average of the equations above with the probabilities of each type of day occurring. So, the Likelihood function becomes

$$\begin{aligned} L(B,S|\theta) &= (1-\alpha)e^{-\varepsilon_b}\frac{{\varepsilon_b}^B}{B!}e^{-\varepsilon_s}\frac{{\varepsilon_s}^S}{S!} \\ &+ \alpha\delta e^{-\varepsilon_b}\frac{{\varepsilon_b}^B}{B!}e^{-(\mu+\varepsilon_s)}\frac{(\mu+\varepsilon_s)^S}{S!} \\ &+ (1-\alpha)\delta e^{-(\mu+\varepsilon_b)}\frac{(\mu+\varepsilon_b)^B}{B!}e^{-\varepsilon_s}\frac{{\varepsilon_s}^S}{S!} \end{aligned} \tag{15}$$

Using the assumption that days are independent, the likelihood of observing the data $D=(B_i,S_i)_{i=1}^{I}$ over I days is the product of the daily likelihoods,

$$L(M|\theta)=\prod_{i=1}^{I}L(\theta|B_i,S_i) \tag{16}$$

To estimate the parameter vector θ, we maximize the likelihood defined in equation (16). This will give us estimates of the parameters in θ and the standard deviations of these estimates will be found using the *delta method*.

Data:

The data were supplied by the Foreign Currency Department of Bank of Israel (BOI). The data consist of daily purchases and sales of currency in the NIS/Dollar market covering six years. The dataset begins on June 2, 2000 and ends on July 3, 2006. After holidays are excluded, the final dataset includes 1488 observations. The data includes spot and swap transactions by customers against the bank but not forward transactions even though there is a perfectly functioning market for that in Israel.

NIS trade against the dollar is the biggest part of total turnover in the NIS foreign exchange market, accounting for 87-90 percent of total turnover over the last five years. In terms of the importance of the market, Israel's foreign exchange activity is comparable to that of countries such as the Czech Republic, Greece and New Zealand. In 2004, average daily turnover was approximately \$1.65 billion for the whole market, just below \$1.68 billion in 2003[8]. In terms of just the spot market, turnover was \$722 million in 2004, slightly lower than the \$748 million figure in 2003[9].

[8] These numbers include swaps as well as spot transactions.
[9] These statistics are taken from BOI's yearly publication analyzing the developments in Israel's currency markets.

The customer order flow in the data is categorized under the following three headings: Foreign Financial Institutions, Other Customers and Domestic Interbank. This is also how the data is broken down in BOI's reports. The foreign financial institutions group includes foreign banks (commercial or investment) with branches in Israel. BOI publications note that foreign financial institutions engage in more speculative activity than domestic customers do and also that they drive the market. On days of high trade turnover, the market share of foreign financials increases significantly compared to their average market share for each year. It is also noted that foreign financial institutions consistently hold larger open foreign exchange positions (both long and short) than the Israeli customers. Because of these descriptions, I hypothesize that foreign financial institutions are the customers who are driving the market (similar to financial customers in Bjonnes et. al. (2005)) and that they are the customers with speculative motives.

The second kind of investor in the market, other customers, are reported to be a heterogeneous group of domestic customers who trade foreign exchange for variety of reasons such as imports and exports, tourism, fund management, investment, and so on. When I inquired further about this group of investors, the head of the Economic Unit at BOI described this group as composed of domestic companies and private customers; by definition, non-financial domestic institutions. For the current analysis, I take this group of other customers to be similar to non-financial customers in Bjonnes et. al. (2005).

Domestic banks are the intraday liquidity providers and Domestic Interbank category covers trades between these banks. Purchase and sale data for the domestic interbank transactions sum to zero most of the time, so I know they are not the ones supplying overnight liquidity in the market.

Table 1 presents descriptive statistics for the order flow in the NIS/Dollar market spanning the whole time series. Moments are quite balanced for both of the customer types. Foreign financial institutions have the top number in terms of buying NIS, whereas other customers have the top number in buying dollars. Standard deviations also seem to be similar with average daily total turnover reaching $570 million. Total turnover average is higher (approximately $630 million) when early years in the data set are left out since those are the years when the market was still maturing. BOI reports indicate that the market matured in 2002 after several years of expansion in response to liberalization.

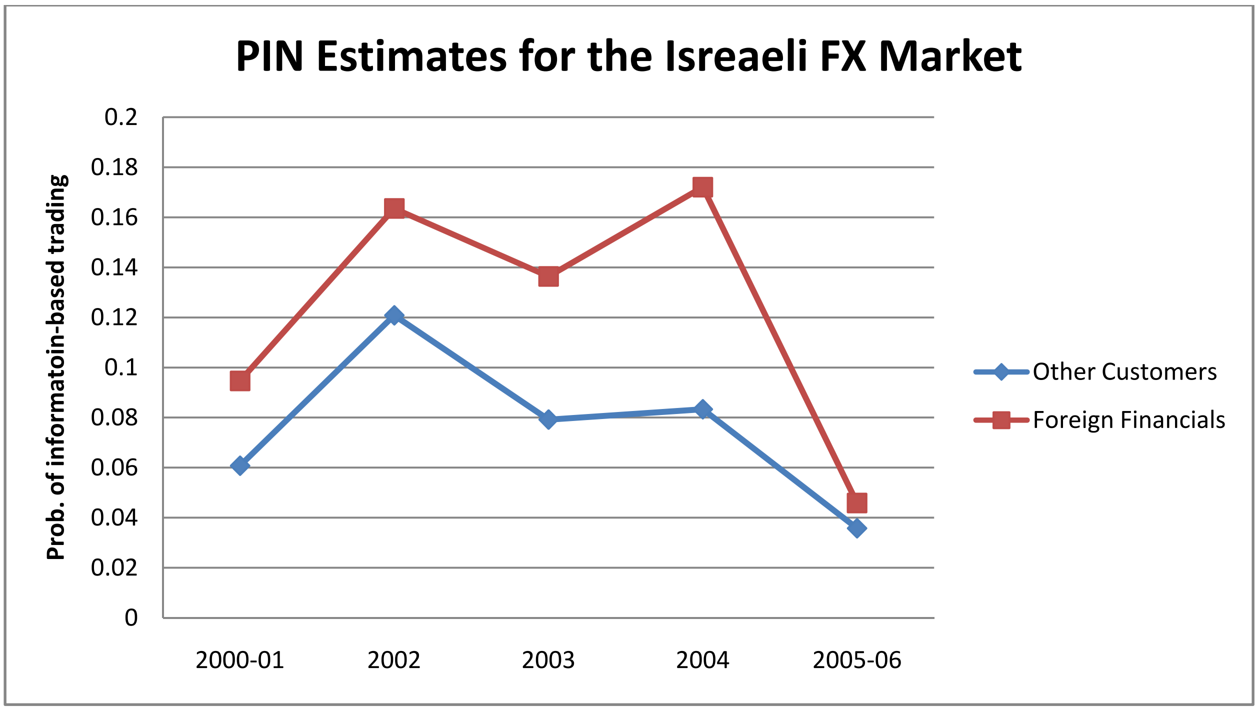
Variables:	**Financial OF**	**Other Cust OF**	**Aggregate OF**	**Total Turnover**
Mean	-3.1815	-2.6736	-5.8552	574.6759
Std. dev.	57.4250	67.0683	62.0811	272.2437
Min.	-335.2210	-262.7980	-256.9210	93.3220
Max.	255.7040	438.3110	307.6140	2164.61

Table1. Descriptive statistics on currency flows in spot market. All values are in millions of Dollars. OF stands for net order flow.

PIN Estimation Preliminary Results:

The dataset used in this study begins on June 2, 2000 and ends on July 3, 2006. We investigate the PIN measure for different spans of time and for both of the groups of investors.

Graph below summarizes the results. Model predicts a higher probability of informed trading for foreign financials than it predicts for other customers. The difference is especially large for the year 2004.



And in order to compare our findings with those in the literature, I present the table below. Marsh and O'Rourke (2005) has a high variance for their findings but when compared to the rest, the probabilities of informed trading our model suggests seems within range of those suggested in the literature.

Literature	**Easley et al. (1996)**	**Easley et al. (2002)**	**Marsh & O'Rourke (2005)**	**Gencay et al. (2007)**	**Gencay et al. (2008)**
PIN Estimates	0.16	0.21	0.09-0.63	0.12	0.11

Next, we make use of delta method to find the standard errors for these estimates and test them to see if they are significant.

References:

Glosten, L. R., and Milgrom, R. P., 1985, "Bid, ask, and transaction prices in a specialist market with heterogeneously informed trader," *Journal of Financial Economics* 14, 71-100.

Easley, D., Hvidjkaer, S., and O'Hara, M., 2002, "Is Information Risk a Determinant of Asset Returns?" *The Journal of Finance,* 57, No. 5, 2185-2221.

Onur, E., 2008, "The Role of Asymmetric Information Among Investors in the Foreign Exchange Market," *International Journal of Finance and Economics*, 13, 368-385.