Information, Liquidity, and Noise [1]

James P. Weston
Jones Graduate School of Management, Rice University
6100 Main Street, Houston, TX 77005
Tele: 713-348-4480
E-mail: westonj@rice.edu

This version: April 2001

[1]Preliminary and Incomplete. Not for quotation. Comments welcome. Address e-mail correspondence James P. Weston at westonj@rice.edu. Thanks To Albert Wang for helpful discussions and comments.

Abstract

This paper presents an empirical analysis of noise trading in the presence of both informed traders and liquidity traders. We model noise traders (following Black (1986)) as agents who trade as though they are informed - when, in fact, they have no information at all. Extending the model of Easley, Kiefer, O'Hara, and Paperman (1996), we distinguish between three types of traders in the market: informed traders, liquidity traders, and noise traders. Our estimates of the relative population of these three trader types suggest that noise traders comprise a significant proportion of the investor population. Further, allowing for the arrival of noise traders significantly reduces estimates of the proportion of informed traders. We find that the probability of informed trading declines from roughly 25 percent to 16 percent with the introduction of noise trading. However, informed trading is still a significant determinant of the bid-ask spread. Additionally, we find that the proportion of noise trading rises - and the falls over volume deciles. Finally, we find that previously documented differences in the probability of informed trading between the Nasdaq and NYSE vanish with the the introduction of noise traders. This is due to the fact that noise trading is significantly higher on the Nasdaq than the NYSE.

Trading volume in many asset markets is curiously high. In foreign exchange, commodity, and equity markets, the total volume of trading generally exceeds what one might expect based simply on portfolio rebalancing needs. Recent work on investor behavior in equity markets points to overconfidence or false perceptions as potential sources for this overly frequent trading.[1] However, investors that fit into this category are hard to classify under standard market microstructure models. They are neither informed traders nor liquidity traders.

Most theoretical market microstructure models classify traders as one of two types - informed or uninformed traders.[2] While informed traders know something about the true value of the asset, uninformed traders enter the market purely for liquidity purposes. However, it is hard to reconcile either the total trading volume or the volatility of trading volume with either of these investor classes. In this paper, we classify a third group of investors as "noise" traders.[3] We model noise traders in the spirit of Black (1986) as investors who believe they have private information but instead have nothing but noise.

To incorporate noise trading, we present an empirical model of the order arrival process following Easley, Kiefer, O'Hara and Paperman (1996) (henceforth EKOP). While EKOP distinguish only between informed and uninformed traders, we separate uninformed traders into liquidity traders and noise traders. We define noise traders as investors who receive a pseudo-signal about the value of the asset. While these traders believe that the signal is correlated with the value of the asset, it is, in fact, pure noise. As Black (1986) states:

[1] See, for example, Odean (2000).

[2] Notable exceptions are Kyle and Wang (1997) and Wang (1999).

[3] Note that many microstructure models use the terms "noise traders", "liquidity traders", and "uninformed traders" synonymously. In this paper, we distinguish between liquidity traders (who have no information and know it) and noise traders who have no information but think that they do. Both of these trader type are, in general, uninformed.

noise are willing to trade even though, from an objective point of view they would be better off not trading. Perhaps they think the noise they are trading on is information. Or perhaps they just like to trade." (page 531)

This concept of noise trading is consistent with the empirical evidence that trading volume seems curiously high and volatile and may also be consistent with the recent rise of day-trading firms that allow individual investors to trade frequently and aggressively, often based on technical trading rules.

The importance of noise trading is not purely academic. Given the strong empirical relationship between the volume of trading and the volatility of asset prices, it is imperative that we understand the composition of volume to discern what factors influence price changes. One reason often cited for the recent rise in equity market volatility is the growth of day-traders and aggressive intra-day trading by individual investors - i.e., noise traders.

Our results show that, in fact, a substantial portion of the trading population consists of noise traders. On average, we find that between 25 and 40 percent of all trades arise from noise traders while only 15 percent are informed. This result is in contrast to the roughly 24 percent informed trading implied by the baseline EKOP model.[4]

In addition, we examine the relationship between trading volume and the relative populations of information, liquidity, and noise traders. Consistent with previous studies, we find the probability of informed trading declines over volume deciles while liquidity trading increases. Interestingly, we find a non-monotonic relationship between volume and the probability of noise trading. Noise trading is small in the lowest volume deciles and increases over medium volume stocks before

[4]While we find a large reduction in the probability of informed trading relative to the EKOP model, we also reconfirm their results in that we continue to find that the probability of informed trading declines with total trading volume and is positively related to the bid-ask spread, even conditional on noise trading.

Finally, we compare the relative populations of trader types by exchange mechanism (NYSE&AMEX vs. Nasdaq). First, we find that noise trading is significantly more prevalent on the Nasdaq than the NYSE or AMEX. Further, we find that previously documented differences in the probability of informed trading between the Nasdaq and NYSE may be driven by noise trading rather than information-based trading. Specifically, we find that after accounting for noise trading, there is only a small difference in the probability of informed trading between the auction and dealers markets. The presence of more noise trading on the Nasdaq may also help explain why prices respond more slowly to information on the Nasdaq relative to the NYSE.

The paper is organized as follows: Section II outlines the model of trade arrival. Section III presents the results of the models. Section IV presents an analysis of the results by volume decile. Section V investigates the relationship between noise trading and the spread. Section VI compares information and noise by exchange type. Section VII concludes.

2 Model

In this section, we outline our model of the trade arrival process. Our set-up is based on the model of Easley, Kiefer, O'Hara and Paperman (EKOP) (1996). First, we present an overview of their model as the baseline case, and then develop extensions to the model that allow for the arrival of noise traders.

The intuition behind the basic order arrival model is straightforward. Each trading day is classified as arising from one of three regimes: a low volume day, a high volume day with more orders to buy than sell, and a high volume day with more sell orders than buys. The order imbalance on high volume days arises from the arrival of informed traders, who know if there is good or bad news about the asset that day.

tion, it does have some drawbacks. For example, on high volume days, trading volume is augmented by informed traders only - and then only on one side of the market. However, on information days there is not only a higher order imbalance, but also a larger amount of *both* buy and sell orders. That is, information days in the EKOP model do not account for the over dispersion in volume between information days and non-information days. When there is a large order imbalance on information days, trading volume is typically abnormally large on *both* the buy and sell side. What drives some days to have large spikes in *both* the number of buy *and* sell orders? We argue that this is driven by noise trading.

2.1 The Baseline EKOP Model

The model presented here follows closely the sequential trade model of Easley, Kiefer, O'Hara and Paperman (EKOP) (1996). The EKOP model relies on the total number of buy and sell trades during a day to identify informed trading. In their model, information events may occur only before the start of trading on each day.[5] The probability of an information event occurring is given by α. Given that an information event has occurred, *bad* news occurs with probability δ while good news has probability $(1-\delta)$. There are two types of stylized traders: informed traders who know the true value of the asset and uninformed traders who trade purely for liquidity purposes. The key feature of this model is that the arrival of these two types of traders are governed by independent Poisson processes. Regardless of information events, the arrival rate of uninformed traders is ε. Informed traders, on the other hand, will arrive to the market only if an information event has occurred, and then only on one side of the market with arrival rate μ. The probability of informed trading is then given by:

$$PI = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}$$

[5]This is in contrast to the Glosten and Harris (1988), where information events may occur at any time.

parameters of the model may be estimates via maximum likelihood. The likelihood of observing B buys and S sells over some time interval T on a given day can therefore be summarized as the sum of three weighted Poisson processes where the weights are given by the probability of an information even, α, and the probability of the bad news, δ:

$$\begin{aligned} L(B,S|\alpha,\delta,\varepsilon,\mu) &= (1-\alpha)\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^B}{B!}\right]\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^S}{S!}\right] \\ &\quad +\alpha(1-\delta)\left[e^{-((\varepsilon+\mu)T)}\frac{((\varepsilon+\mu)T)^B}{B!}\right]\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^S}{S!}\right] \\ &\quad +\alpha\delta\left[e^{-(\varepsilon T)}\frac{(\varepsilon T)^B}{B!}\right]\left[e^{-((\varepsilon+\mu)T)}\frac{((\varepsilon+\mu)T)^S}{S!}\right] \end{aligned}$$

Assuming that the days are independent, the likelihood of observing a sequence of buys and sells $X = (B_i, S_i)_{i=1}^N$ over N days is simply:

$$L(X|\alpha,\delta,\varepsilon,\mu)_{baseline(EKOP)} = \prod_{i=1}^{N} = L_i(B_i, S_i|\alpha,\delta,\varepsilon,\mu) \quad (1)$$

2.2 Noise model 1: Incorporating Partially informed Noise traders.

In this subsection, we incorporate the arrival of noise traders in the EKOP model. Noise traders here are defined as traders who get a clean signal of whether each day is an information day or not, i.e., α. These traders also receive a pseudo-signal of whether the information is good news or bad news. However, unlike informed traders, noise traders receive a signals of good and bad news with equal probability, regardless of the true value of δ. In this sense, the noise traders know that there is information in the market, but incorrectly assume that they have an accurate signal of the quality of that news.

The structure of the model for the arrival of noise traders implies that, on an information day, total volume will be augmented by the arrival of noise traders on both the buy and sell side. That is, total volume on an information day will consist of liquidity traders (arriving on the buy and sell like liquidity traders, arrive on both the buy and sell sides of the market. Like liquidity traders and informed traders, we assume that noise traders arrive to the market by an independent Poisson process with arrival rate η.

Thus, on a bad-event day, the likelihood of observing a sequence of B buy orders and S sell orders is simply

$$\left[e^{-(\varepsilon+\eta)}\frac{(\varepsilon+\eta)^B}{B!}\right]\left[e^{-(\varepsilon+\mu+\eta)}\frac{(\varepsilon+\mu+\eta)^S}{S!}\right] \tag{2}$$

On a good-event day, the arrival rate is :

$$\left[e^{-(\varepsilon+\mu+\eta)}\frac{(\varepsilon+\mu+\eta)^B}{B!}\right]\left[e^{-(\varepsilon+\mu+\eta)}\frac{(\varepsilon+\eta)^S}{S!}\right] \tag{3}$$

On a no-event day, neither informed traders nor noise traders will arrive to the market. Thus, the likelihood of observing B buys and S sells is

$$\left[e^{-(\varepsilon)}\frac{(\varepsilon)^B}{B!}\right]\left[e^{-(\varepsilon T)}\frac{(\varepsilon)^S}{S!}\right] \tag{4}$$

The introduction of noise traders into the EKOP model does not alter the information parameters, α and δ. As a result, the probability of observing the three arrival rate regimes do not change from the baseline model. The likelihood of observing a sequence of buys and sells $X = (B_i, S_i)_{i=1}^N$ over N days is simply given by

$$\begin{aligned} L(B,S|\alpha,\delta,\varepsilon,\mu,\eta) &= (1-\alpha)\left[e^{-(\varepsilon)}\frac{(\varepsilon)^B}{B!}\right]\left[e^{-(\varepsilon T)}\frac{(\varepsilon)^S}{S!}\right] \\ &\quad +\alpha(1-\delta)\left[e^{-(\varepsilon+\mu+\eta)}\frac{(\varepsilon+\mu+\eta)^B}{B!}\right]\left[e^{-(\varepsilon+\mu+\eta)}\frac{(\varepsilon+\eta)^S}{S!}\right] \\ &\quad +\alpha\delta\left[e^{-(\varepsilon+\eta)}\frac{(\varepsilon+\eta)^B}{B!}\right]\left[e^{-(\varepsilon+\mu+\eta)}\frac{(\varepsilon+\mu+\eta)^S}{S!}\right] \end{aligned} \tag{5}$$

Again, sssuming that the days are independent, the likelihood of observing a sequence of buys

$$L(X|\alpha,\delta,\varepsilon,\mu)_{Noise_1} = \prod_{i=1}^{N} = L_i(B_i,S_i|\alpha,\delta,\varepsilon,\mu,\eta) \tag{6}$$

This model is equivalent to the EKOP model except for the addition of the parameter η on information days. Note that if $\eta = 0$, equation 5 collapses to equation 1, the baseline EKOP model. Given an estimate of the parameter vector $[\alpha, \delta, \varepsilon, \mu, \eta]$ we can construct the probability of informed trading as :

$$PI = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon + \alpha 2\eta} \tag{7}$$

Similarly, we can construct the probability of noise trading as:

$$PN = \frac{\alpha 2\eta}{\alpha\mu + 2\varepsilon + \alpha 2\eta} \tag{8}$$

2.3 Noise model 2: Incorporating uninformed Noise traders

In the previous sub-section, we made the assumption that noise traders received a clean signal of α but an infinitely noisy signal of δ. In this sub-section, we relax the assumption that noise traders receive a clean signal of α. Instead, we assume some probability, κ, that noise traders receive a correct signal of α. For example, on a bad-event day there is a κ probability that noise traders will receive that correct signal and arrive to the market. Conversely, on a no-event day, there is a $(1-\kappa)$ probability that noise traders will arrive to the market, incorrectly assuming that it is an information day. This framework is slightly more complicated than the previous model and gives rise to six different arrival regimes

Regime	Event	Arrival of Buys	Arrival of Sells	Probability
1	No-event	ε	ε	$\kappa(1-\alpha)$
2	No-event	$\varepsilon+\eta$	$\varepsilon+\eta$	$(1-\kappa)(1-\alpha)$
3	Good-event	$\varepsilon+\mu$	ε	$(1-\kappa)(\alpha)(1-\delta)$
4	Good-event	$\varepsilon+\eta+\mu$	$\varepsilon+\eta$	$(\kappa)(\alpha)(1-\delta)$
5	Bad-event	ε	$\varepsilon+\mu$	$\kappa(\alpha)(\delta)$
6	Bad-event	$\varepsilon+\eta$	$\varepsilon+\eta+\mu$	$(1-\kappa)(\alpha)(\delta)$

write the likelihood as in equation 5 where each of the six Poisson arrival regimes are weighted by the state probabilities. Again, assuming that the days are independent, the likelihood of observing a sequence of buys and sells $X = (B_i, S_i)_{i=1}^{N}$ over N days is simply:

$$L(X|\alpha, \delta, \varepsilon, \mu)_{Noise_2} = \prod_{i=1}^{N} = L_i(B_i, S_i|\alpha, \delta, \varepsilon, \mu, \eta, \kappa) \tag{9}$$

This model is also similar to the baseline model with the addition of the parameters κ and η. Note that if $\kappa = 1$ or $\eta = 0$ then the model collapses to the baseline EKOP model. Similarly if $\kappa = 0$ then the model collapses to the simple noise case. Given our estimates of both the restricted and unrestricted models, we can test the performance of the models, relative to the baseline. As in the simple noise trading model, given an estimate of the parameter vector $[\alpha, \delta, \kappa, \varepsilon, \mu, \eta]$ we can construct the probability of informed trading as :

$$PI = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon + (1-\alpha)(1-\kappa)2\eta + \alpha\kappa 2\eta} \tag{10}$$

Again, the probability of noise trading can be constructed as:

$$PN = \frac{(1-\alpha)(1-\kappa)2\eta + \alpha\kappa 2\eta}{\alpha\mu + 2\varepsilon + (1-\alpha)(1-\kappa)2\eta + \alpha\kappa 2\eta}$$

3 Data

We use data for all common, domestic stocks listed securities on the NYSE, AMEX or Nasdaq during the first quarter of 1998. From the universe of all stocks, we select only those firms that we can match from the CRSP and TAQ database based on the ticker symbol. This reduces our sample to 8,101 stocks. We then limit our sample to firms that trade on at least 25 days during the quarter. This filter reduces our sample to 5,588 stocks. We then estimate the baseline-EKOP model and our two noise trading models for all stocks in our sample. Given the computational intensity of this exercise, we also eliminate firms for which the likelihood function did not converge. This

2,297 from the NYSE and only 174 from the Amex. While the likelihood function is generally well-behaved, many of the firms for which the likelihood function did not converge are very large firms with a highly volatile number of daily transactions. For these firms, it was difficult to evaluate the likelihood function on certain days under different starting values. Note that this biases our sample of firms towards mid-cap and small-cap stocks.

Table 1 provides an overview of the firms in our sample. The stocks in our sample have an average price near $18 with an average monthly trading volume of 209,000 shares. There is a considerable amount of dispersion in the sample over the number of trades and share turnover. The average firm in our sample trades 93 times per day while roughly 95% of the firms trade more than 10 times per day. There is also substantial variation in the trading costs for our sample. The average quoted bid-ask spread generally varies between 7 cents and 43 cents which represents between 0.1 percent and 8 percent of the price.

In order to estimates the number of buy and sell orders for each stock-day, we use a modified version of the Lee and Ready (1993) algorithm. Trades are matched to the relevant quotes using a 16-second delay as suggested by Blume and Goldstein (1998). That is, for each trade, we take the most recent bid and ask quote that was posted at least 16 seconds before the trade was executed. This procedure accounts for the fact that quotes and trades are recorded on separate systems and may be systematically non-synchronous (Lee and Ready (1991)). To the extent that trades take longer than 16 seconds to be reported, we introduce measurement error into estimates that rely on both trade prices and quotes (e.g., effective spreads and our trade indicator variables). Lee and Ready suggest using a five second delay to correct for the different reporting times. While the five-second rule is the most commonly used, Hasbrouck (1993) reports a median delay of 14 seconds. If the delay is too long however, then we could err on the side of matching trades to stale quotes. Nevertheless, these concerns are partially mitigated by the fact that averages are computed bias using the 16-second algorithm. If a trade occurs inside the bid-ask spread, we use a tick-test to classify the trade. Having classified all trades as buys or sells, we then sum the total number of buy and sell orders for each stock-day over the 65 trading days in the sample.

4 Results

In this section, we present the results of our estimation of equations (1), (5), and (9) for our sample of 4,906 stocks. To estimate the parameters of the models, we use constrained maximum likelihood, using a series of different starting values. We retain only those observations for which we achieve a stable convergence to the optimum. The Poisson arrival rates, ε, μ, and η, are constrained to the interval $[0, \infty)$ while the probabilities, α, δ, and κ are constrained to the interval [0,1].

Table 2 presents the results of our estimation of the parameters along with the probability of informed trading, PI. Standard errors for the PI are constructed via the delta method. Overall, we find that our estimates of $\alpha, \delta, \varepsilon$, and μ are consistent with O'Hara (1999). However, we find that the introduction of noise trading into the model does significantly affect the parameters. First, both noise models find, on average, a higher value of α than under the baseline model. This results is consistent with the greater flexibility in modeling over dispersion in the total number of buys *and* sells on information days. However, we find no statistical difference in the probability of bad news, δ, between the three models. This is re-assuring since there is no reason to expect the introduction of noise traders to affect this parameter.

The main effect of the noise traders can be seen in comparing our estimates of ε and μ across the three models. We find that both ε and μ are significantly higher (using a K-Wallis test and Man-Whitney tests) in the baseline EKOP model than in either of the noise trading models. In fact, we find over a 40 percent decline, on average, in the intensity of the arrival of informed traders, noise traders.

Noise trading has a significant effect on our estimate of the probability of informed trading. Comparing the last three rows of Table 2, we find that the probability of informed trading declines by roughly 9 percentage points (45%) with the introduction of noise trading. These results point to the importance of noise trading in equity markets and demonstrate that much of the over dispersion in volume may be driven by the population of noise traders.

While Table 2 presents the results for the parameters of the model, we cannot say whether the noise-trading models are any better at describing the pattern of trading volume for the stocks in our sample. Since the baseline EKOP model is nested within our two noise-trading models, we can test whether the models perform better than the baseline model by constructing a likelihood ratio statistic for each of the models. Table 3 presents the results of the likelihood ratio test for all of the stocks in our sample. The null hypothesis is that the unrestricted models (the noise trading models) are no better at maximizing the likelihood function than the restricted model (the baseline-EKOP model). Table 3 shows that, for the vast majority of stocks in our sample, we cannot reject the presence of noise trading.

5 Results by Volume Decile

While the results presented in section 4 show that noise traders are a substantial portion of the trading population, the results hide some of the differences across volume deciles. Easley, Kiefer, O'Hara and Paperman (1996) find that the probability of informed trading declines with total trading volume. Our results continue to support this hypothesis in the presence of noise trading. Table 4 presents estimates of our parameters over three different volume deciles. We report results only for deciles one, five and ten for simplicity. The results clearly show that, while the arrival informed trading declines. That is, the increase in uninformed trading increases over the volume deciles at a greater rate than the increase in informed traders.

Figures 1,2, and 3 show the probabilities of informed, noise, and liquidity trading, respectively, over the ten volume deciles. The results for informed and liquidity trading are consistent with the fact that informed trading is more concentrated in low-volume stocks. However, the pattern in the probability of noise trading over the volume deciles show a somewhat more complex relationship. For both models of noise trading, the probability of noise trading first increases, then decreases over the volume deciles, suggesting the noise trading is most concentrated in mid-volume stocks and lowest in the very high volume stocks. While the probability of noise trading may be lowest in the highest volume stocks, noise traders still constitute between 15 and 20 percent of the total trading population in the largest volume decile.

6 Relationship with the Spread

In this section we test whether noise trading affects the relationship between informed trading and the bid-ask spread. Easley, Keifer, O'Hara and Paperman (1996) show that the spread may be approximated as:

$$Spread = \beta_1 * \Pr(Informed) * \Pr ice \tag{11}$$

Incorporating the effect of trading volume on the inventory costs faced by intermediaries, we estimate the following regression equation:

$$\begin{aligned} \ln(Spread) &= \beta_0 + \beta_1 * \ln(\Pr(Informed)) + \beta_2 \ln(\Pr ice) \\ &\quad + \beta_3 * \ln(Turnover) + \beta_4 Nasdaq + \beta_5 AMEX + \xi \end{aligned} \tag{12}$$

We estimate the same regression for our three models. Table 5 presents the results of the three regressions. Overall, we find consistent results that turnover is negatively related to the spread models of market maker pricing and with institutional features of the equity markets. In addition, the coefficients on both the AMEX and Nasdaq exchange dummy variables are positive and significant, suggesting that spreads are smallest on the NYSE. This results is also consistent with prior research.[6]

Our results are roughly consistent with the hypothesis that informed trading is positively associated with the bid-ask spread. For the baseline model, we find a positive and significant coefficient on the probability of informed trading suggesting that a one percent increase in the PI is associated with a 0.44 percent increase in the spread. After incorporating simple noise trading into the model, we continue to find a positive and significant effect of information on spreads, though the effect is somewhat smaller (0.18 compared to 0.44 for the baseline case). However, in the extended noise trading model, we find no significant relationship between the spread and the PI.

7 Results by Exchange Type

In sections 5 and 6 we show that a substantial proportion of the investor population is comprised of noise traders. A number of recent studies have argued that there may be systematic differences in the information structure of different exchange mechanisms. For example, Hiedle and Huang (2000) examine a sample of firms that switch exchange listing and argue that the probability of informed trading (using the baseline EKOP model) is higher on the Nasdaq than on the NYSE or AMEX. The reason for this, they argue, may be that the Nasdaq is an anonymous quote-driven market whereas on the NYSE and AMEX, the specialist knows the identity of each trader. As a result, informed traders, on the margin, have more of an incentive to arrive on the Nasdaq than on the NYSE. In this section, we compare our estimates of the relative trader-type populations across the different market mechanisms.

[6]See, for example, Huang and Stoll (1995).

with Z-statistics[7] of the difference between the exchanges. Consistent will previous research, we find that the baseline EKOP model indicates a higher proportion of informed trading on the Nasdaq, relative to the NYSE/AMEX sample. However, once noise traders are introduced into the model, we find that the PI measure is actually smaller for our sample of Nasdaq stocks.

The driving force of this results stems from the difference in our estimates of the probability of noise trading across the two exchange mechanisms. For both noise trading models, we find that the probability of noise trading is over 50% larger on the Nasdaq than on the NYSE/AMEX. These results suggest that investors who trade Nasdaq stocks are more likely to receive pseudo-signals as opposed to clean signals. That is, we find that the baseline EKOP model incorrectly labels many noise traders as informed traders, based simply on the relatively greater over dispersion of trading volume on that exchange. Once the over dispersion in volume is incorporated into the model, the difference in informed trading reverses.

The result that the Nasdaq is more heavily populated by noise traders is consistent with the type of stocks where noise trading is more likely to occur - in assets where clean signals may be hard to identify, namely high-technology firms that are more represented on the Nasdaq. The result is also consistent with the trading preferences of day-traders.

8 Conclusion

This paper presents a empirical investigation of the role of noise traders in U.S. equity markets. We find that a substantial proportion of the trading population is comprised of noise traders, i.e. investors who trade on noise as if it were information. On average, we find that between 25 and 40 percent of the orders for our sample of stocks are from noise traders. Further, we find

[7] The Z-statistics are based on Mann-Whitney pair wise comparisons of the two samples. The null hypothesis is that the two samples are drawn from the same distribution. The test statistic is asymptotically normal.

information-based trading.

In addition, we find a non-linear relation between trading volume and the population of noise traders suggesting the noise trading is mode heavily concentrated in mid-volume stocks. Further, while we find a large reduction in the probability of informed trading relative to the baseline model, we continue to find that informed trading is positively related to the bid-ask spread. finally, our results show that noise trading is far more common on the Nasdaq than on the NYSE/AMEX.

[1] Black, Fischer, 1986, Noise, *Journal of Finance* 49, 529-543.

[2] Blume, Marshall E., and Michael A. Goldstein, 1997, Quotes, Order Flow, and Price Discovery, *Journal of Finance* 52, 221-44.

[3] Easley, David, and Nicholas M. Kiefer, 1996, Liquidity, information, and infrequently traded stocks, *Journal of Finance* 51, 1405-1436.

[4] Garfinkel, Jon A., and M. Nimalendran, 1998, Market Structure and Trader Anonymity: An Analysis of Insider Trading, Working Paper.

[5] Glosten, Lawrence R., and Robert Milgrom, 1985, Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71-100.

[6] Hasbrouck, Joel, and George Sofianos, 1993, New York Stock Exchange systems and trading procedures, NYSE Working paper 93-01.

[7] Heidle, Hans, and Roger D. Huang, 1999, Information-Based Trading in Dealer and Auction Markets: An Analysis of Exchange Listings, Working paper, Vanderbilt University.

[8] Huang, Roger D., and Hans Stoll, 1996, Dealer vs. auction markets: A paired comparison of execution costs on Nasdaq and the NYSE, *Journal of Financial Economics* 41, 313-357.

[9] Kyle, Albert S., and F. Albert Wang, 1997, Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?, *Journal of Finance* 52, 2073-2090.

[10] Lee, Charles M., and Mark J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733-46.

[11] Odean, Terrance, 2000, Do investors trade too much ?, forthcoming *American Economic Review*.

[12] O'Hara, Maureen, 1999, Making Market Microstructure Matter, *Financial Management* 28, 83-90.

[13] Trueman, Brett, 1988, A Theory of Noise Trading in Securities Markets, *Journal of Finance* 1988, 83-95.

[14] Wang, F. Albert, 2000, Speculative Noise Trading in Financial Markets, Working paper, Rice University.

Table 1:

Summary Statistics

This table presents descriptive statistics for our sample of firms.

Descriptive Variables	Obs	Mean	Std.	5%	median	95%
Price	4906	18.28	20.79	1.32	13.29	50.77
Volume	4906	209	510	11	68	799
Number of Trades	4906	93.52	161.34	9.56	38.66	356.31
Share Turnover	4906	5.44	8.72	0.92	3.83	14.49
Volatility	4906	0.05	0.03	0.01	0.04	0.10
Quoted Spread	4906	0.19	0.19	0.07	0.15	0.43
Relative Spread	4906	0.03	0.03	0.00	0.02	0.08

Table 2:

Parameter estimates

This table presents descriptive statistics for the parameter estimates of the three models.

Parameter	Obs	Mean	Std.	5%	median	95%
α_{EKOP}	4754	0.333	0.127	0.141	0.324	0.554
α_{Noise_1}	4902	0.367	0.162	0.125	0.357	0.658
α_{Noise_2}	4837	0.401	0.164	0.140	0.395	0.673
δ_{EKOP}	4754	0.445	0.189	0.144	0.436	0.765
δ_{Noise_1}	4902	0.442	0.187	0.146	0.431	0.760
δ_{Noise_2}	4837	0.470	0.199	0.142	0.472	0.802
ε_{EKOP}	4754	36.5	67.7	3.52	14.1	140.2
ε_{Noise_1}	4902	33.4	63.1	2.94	12.2	130.4
ε_{Noise_2}	4837	27.8	55.0	2.51	9.4	111.0
μ_{EKOP}	4754	46.2	54.0	6.81	27.0	148.9
μ_{Noise_1}	4902	26.6	28.5	4.85	17.2	76.7
μ_{Noise_2}	4837	30.9	34.4	5.53	20.0	91.4
η_{EKOP}	4902	29.6	50.7	1.42	11.6	121.2
η_{Noise_1}	4837	39.9	56.7	2.70	18.6	149.1
κ_{Noise_2}	4837	0.556	0.165	0.304	0.550	0.836
$\Pr(Inf)_{EKOP}$	4754	0.239	0.068	0.114	0.214	0.333
$\Pr(Inf)_{Noise\_2}$	4902	0.150	0.074	0.051	0.139	0.284
$\Pr(Inf)_{Noise\_2}$	4837	0.152	0.070	0.057	0.140	0.282

Table 3:

Likelihood-Ratio Test Statistics

This table presents likelihood-ratio test statistics for the three models.

Hypothesis	Average LR-Statistic	Number significant
$H_0^1 : \eta = 0$	244.5	4,335 (91%)
$H_0^2 : \eta = 0 \; ; \kappa = 0$	414.8	4,225 (88%)
$H_0^3 : \kappa = 0$	182.1	3,971 (82%)

Table 4:

Average parameter estimates by volume decile

This table presents average estimates of the parameters by volume deciles.

Parameter	1st Decile	5th decile	10th Decile
α_{EKOP}	0.318	0.333	0.393
α_{Noise_1}	0.357	0.378	0.433
α_{Noise_2}	0.319	0.396	0.500
δ_{EKOP}	0.471	0.472	0.294
δ_{Noise_1}	0.472	0.471	0.296
δ_{Noise_2}	0.473	0.498	0.328
ε_{EKOP}	4.64	13.74	173.53
ε_{Noise_1}	3.84	11.10	153.67
ε_{Noise_2}	3.23	8.93	135.49
μ_{EKOP}	11.62	28.99	131.20
μ_{Noise_1}	7.80	16.22	72.46
μ_{Noise_2}	9.35	19.18	91.50
η_{EKOP}	4.23	13.84	73.29
η_{Noise_1}	6.34	21.79	108.93
κ_{Noise_2}	0.674	0.545	0.468
$\Pr(Inf)_{EKOP}$	0.260	0.234	0.135
$\Pr(Inf)_{Noise\_2}$	0.205	0.166	0.090
$\Pr(Inf)_{Noise\_2}$	0.172	0.166	0.118

Table 5:

Regression Results

Dependent variable : ln(Quoted Spread)			
$\ln(Pr(Inf)_{EKOP})$	0.440	–	–
	(0.033)		
$\ln(Pr(Inf)_{Noise\_1})$	–	0.183	–
		(0.016)	
$\ln(Pr(Inf)_{Noise\_2})$	–	–	-0.001
			(0.017)
$\ln(Price)$	0.345	0.325	0.292
	(0.009)	(0.009)	(0.009)
$\ln(Turnover)$	-0.109	-0.106	-0.152
	(0.011)	(0.011)	(0.010)
Amex dummy	0.282	0.317	0.306
	(0.041)	(0.041)	(0.041)
Nasdaq dummy	0.334	0.423	0.404
	(0.019)	(0.018)	(0.019)
Constant	-2.020	-2.343	-2.565
	(0.048)	(0.035)	(0.038)
N	4754	4902	4837
R2	0.313	0.300	0.278

Table 6:

Parameter estimates: By exchange

Variable	NYSE/AMEX	Nasdaq	Difference	Z-Stat
$\Pr(Inf)_{EKOP\_2}$	0.195	0.250	-0.055	-23.06
$\Pr(Inf)_{Noise\_1}$	0.154	0.146	0.008	4.36
$\Pr(Inf)_{Noise\_2}$	0.167	0.137	0.030	15.50
$\Pr(Liquidity)_{EKOP\_2}$	0.805	0.760	0.045	23.03
$\Pr(Liquidity)_{Noise\_1}$	0.700	0.607	0.093	30.54
$\Pr(Liquidity)_{Noise\_2}$	0.535	0.355	0.180	44.07
$\Pr(Noise)_{Noise\_1}$	0.146	0.247	-0.101	-42.11
$\Pr(Noise)_{Noise\_2}$	0.297	0.508	-0.210	-46.40

Conditional on $\alpha = 1$				
$\Pr(Inf)_{EKOP\_2}$	0.399	0.533	-0.134	-37.75
$\Pr(Inf)_{Noise\_1}$	0.259	0.247	0.012	4.14
$\Pr(Inf)_{Noise\_2}$	0.322	0.314	0.008	1.24

Figure 1

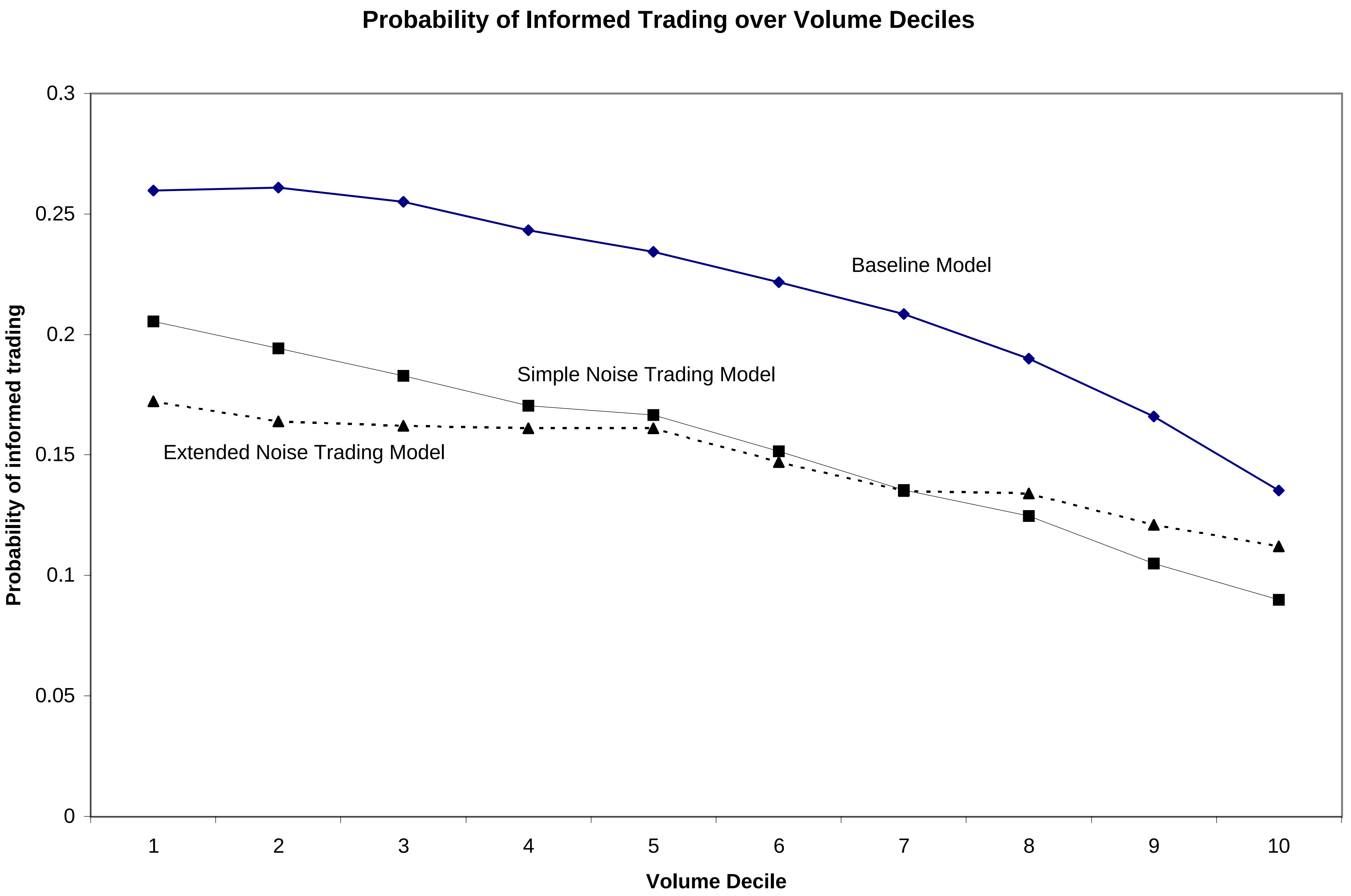
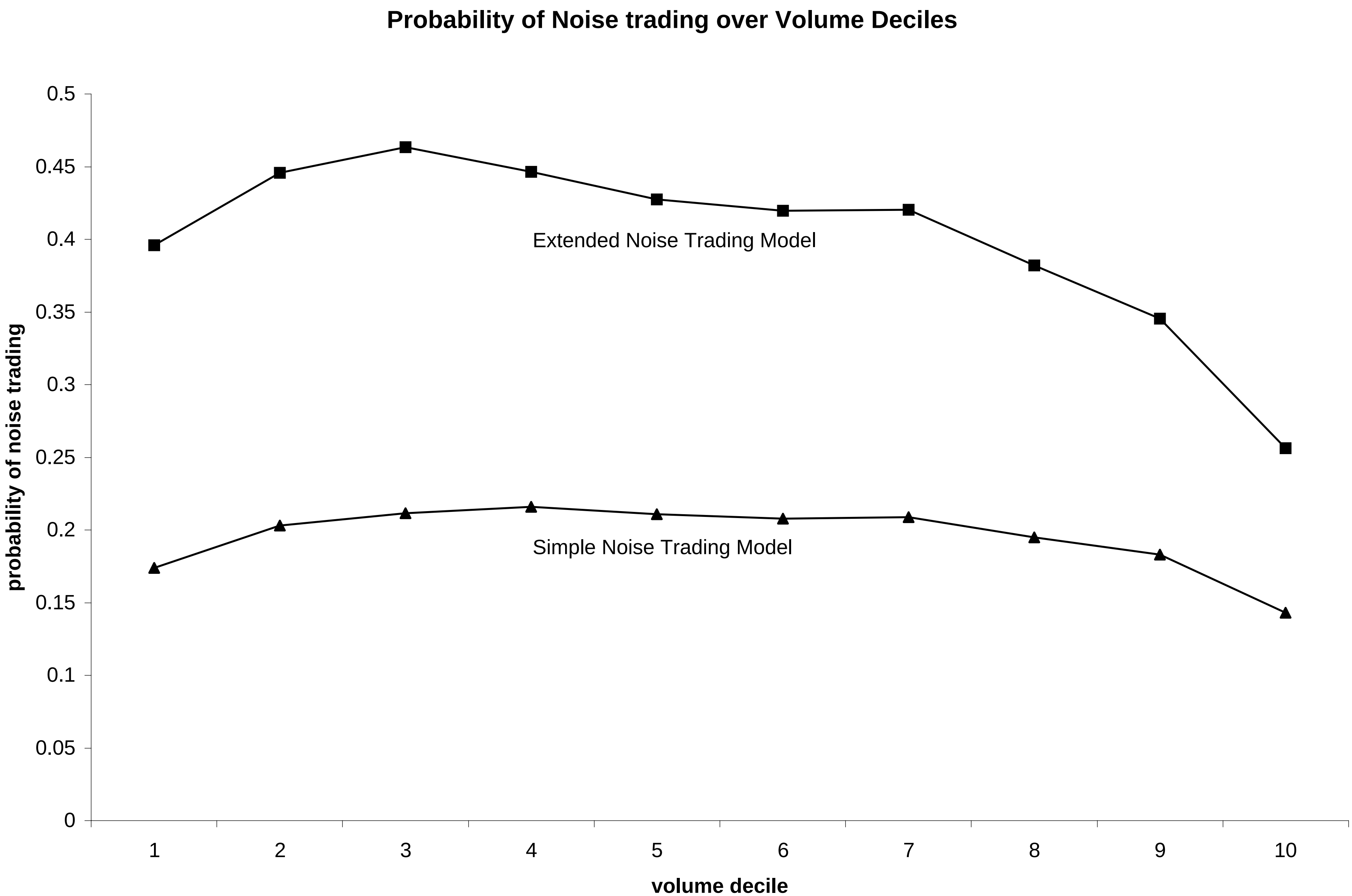
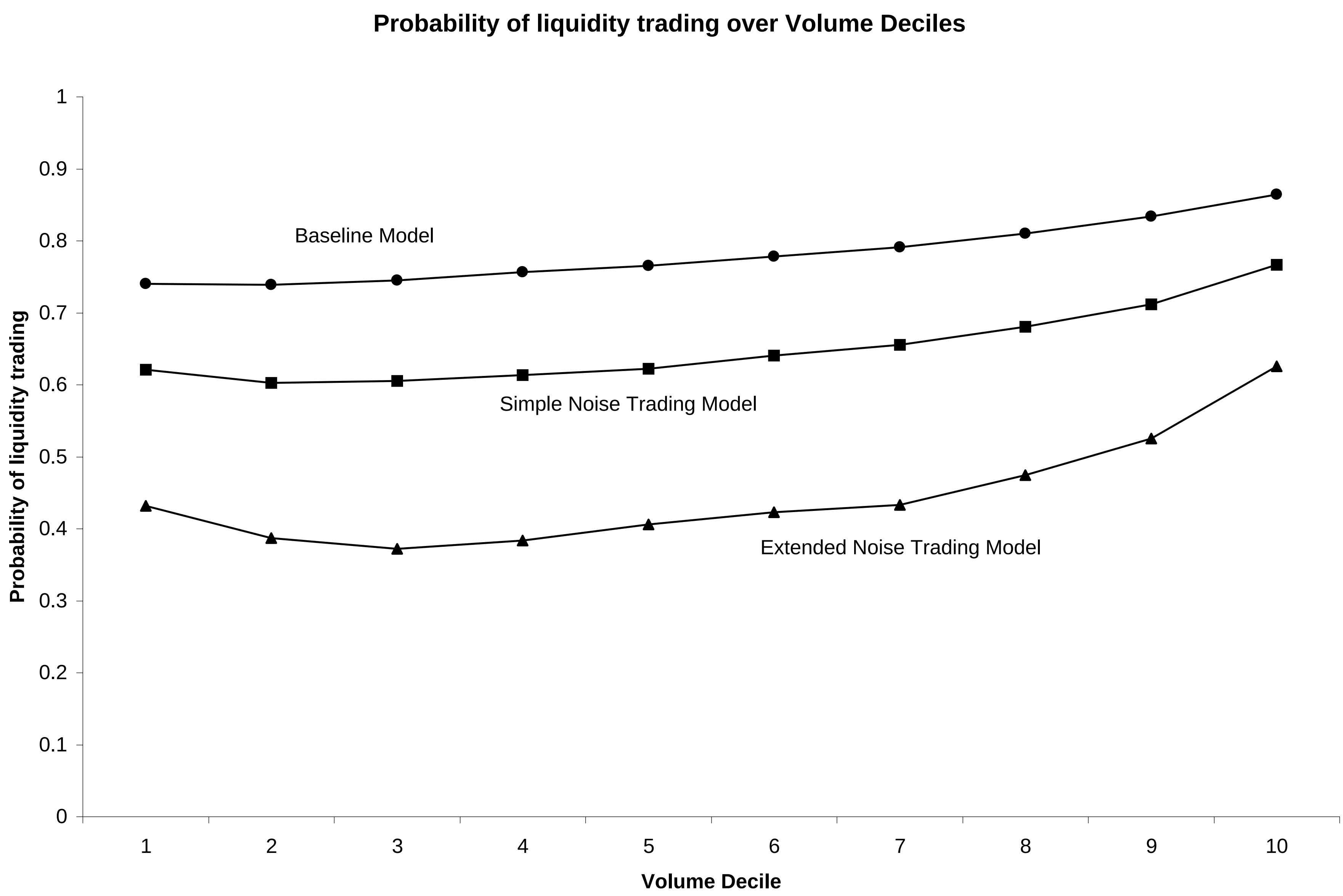

Probability of Informed Trading over Volume Deciles
Probability of informed trading
Volume Decile
0.3
0.25
0.2
0.15
0.1
0.05
0
1
2
3
4
5
6
7
8
9
10
Baseline Model
Simple Noise Trading Model
Extended Noise Trading Model

Figure 2



Probability of Noise trading over Volume Deciles
probability of noise trading
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Extended Noise Trading Model
Simple Noise Trading Model
1
2
3
4
5
6
7
8
9
10
volume decile

Figure 3



Probability of liquidity trading over Volume Deciles
Probability of liquidity trading
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Baseline Model
Simple Noise Trading Model
Extended Noise Trading Model
1
2
3
4
5
6
7
8
9
10
Volume Decile