Chapter Three

A MIXING MODEL WITH ZERO INTELLIGENCE TRADERS

1 Introduction

The characterization of trader behavior may be the most difficult task for the analysis of markets. The trading institution can be characterized as having a definite structure, a set of rules, and a history; while the actions of trader participants might be thought of as based on mental abilities. Gode & Sunder (1993) examine the zero intelligence (ZI) benchmark which models traders' activity as if they followed a simple non-strategic algorithm. The algorithm instructs traders to submit random bids and asks as long as these orders would result in a profitable trade if a transaction were to take place. Easley & Ledyard (1993) show that trading model populated with a variation of these ZI traders achieves surprising efficiency allowing the traders to extract a high percentage of the potential trade surplus in a double auction market. More complex theoretical models include those of Wilson (1987) and Friedman (1991). These models require greater computational skills on the part of traders. The volume edited by Friedman and Rust (1993) includes an overview of theoretical and experimental work on this subject.

The results of simulations by Gode and Sunder demonstrate that ZI traders can serve as a useful benchmark for trader behavior in double auction markets under various institutional rules governing trade execution. While this perspective on trader behavior has been useful to the researcher, strategic traders themselves may gain a better perspective by assuming other traders act as ZI traders. The market institution considered is a double auction market with common values where traders are asymmetrically informed. We assume only one trader approaches the market strategically to avoid strategic recursions. It is then shown that a single strategic trader could use the assumption of ZI to identify the behavior of informed and uninformed traders in this market. Once the trader types are identified, the strategic trader could properly weight the actions of the informed trader and uninformed trader. By categorizing actions according to trader type, a strategic trader can make better use of the available information in the market.

In a market considered, the informed group of traders receives a more precise price

signal than an uninformed group. Each set of signals is drawn from a Normal distribution around the true worth of the asset. The individual signals are private information while the structure of the signal generation mechanism is common information. If each trader bases a market order on private information, it can be shown that the resulting prices alone will not reveal the private information of traders due to the non-linearity of the signal and price relation. This is due to the fact that signals are drawn from two unique distributions, and prices do not identify the distribution of the associated signals.

The key to the problem is the observation that a mix of Normal distributions results in a distribution which is not Normally distributed. This observation leads to interesting results. If the observed market price were based entirely on price signals which were drawn from two Normal distributions rather than a single Normal distribution, the result would be that price would not be Normally distributed. It follows that conventional forecasting techniques predicting the conditional expected mean value of price given past prices cannot be used because the model does not have an error structure which is Normal. Estimates from misspecifying the model as having a Normal error structure will be biased (see also Chapter 1). Classifying prices according to trader type under the ZI assumption is equivalent to identifying the distribution from which the price signal is drawn. Knowing the components of the mix of Normals allows a trader to avoid the non-linear estimation problem, and allows unbiased predictions of the true asset worth.

In what follows, the empirical work on the distribution of prices and price changes is reviewed. Models which assume a mixed Normal structure are then discussed. A technique for estimating a mixed model through a variation of maximum likelihood estimation, the estimation maximum likelihood technique, is then introduced. This is followed by the estimation of a mixture model under the zero intelligence assumption. The data used is from the experimental sessions discussed in Chapter 2. The last section concludes.

2 Empirical Studies of the Distribution of Prices

The characterization of the distribution of security returns has been of interest to researchers back to at least to the time of Bachelier (1900). Empirical studies of historical security returns data show convincingly that returns are not Normally distributed. Clark (1973) is an important reference. An example of extensive empirical work is Friedman and Vandersteel (1980) where daily spot foreign exchange prices from 1973 to 1979 are characterized. It was found that while returns tend to be symmetric, the kurtosis coefficient of these returns was much larger than what would be expected for a Normal distribution. Taylor (1985) describes data for 15 US stocks, a foreign stock index, 6 metals, the dollar/sterling spot rate, and futures on commodities and exchange rates. For the period examined, often exceeding ten years of daily observations, it is found again that the coefficient of kurtosis for these returns exceed the value of 3, the level at which these returns could be considered Normally distributed.[1]

A second feature of security returns is that volume and price changes are positively related. (see, e.g., survey by Karpoff (1987)) Volume has been considered to be the driving process in price variability. Clark (1973) describes variance of prices and volume as having a curvilinear (nonlinear) relation, and demonstrates how this might be modeled as a subordinated stochastic process. The base process is the random arrival of new information. The secondary process is the observed price sequence. Using the price series alone, price changes are found not to be Normally distributed. However, using volume to better estimate the arrival of new information, kurtosis is reduced to levels comparable to those expected under Normal distributions. Volume is also used to explain price variations by Blume, Easley & O'Hara (1994) where the absolute value of prices changes and volume are positively related. Conrad, Hameed, and Niden (1994) find that abnormal volume is related to return autocovariances in an empirical study of equities. Wang (1994) also finds a relation between price changes and volume in his recent

[1] Explanations for non-Normal returns have included the description of returns as subordinated processes or as a mixture of distributions. See Titterington, Smith, Makov for explanation of non-Normality of mixtures of Normals as a corollary of identifiability. p. 39.

theoretical model.

A third feature of security prices is the observed variance heterogeneity of returns over time. The variance of observed price changes has been modeled as having been generated by a mixture of Normals. Observed trading volume has been used to identify this mixture by Tauchen and Pitts (1983), where a bivariate Normal mixture model is used to describe the relation of volume with the variance of price changes. This allows the determination of a conditional distribution of price changes given volume. Their model is then used to study the relation of volume and open interest to the volatility of prices in 3-month T-bill futures contracts.

3 The Estimation Maximum Likelihood Algorithm

3.1 Background

The estimation maximum likelihood (EM) algorithm as a method of solution to finite mixture densities is related to Fisher's method of scoring (see, e.g., Fisher (1935)). The application of the EM algorithm to estimate the components of a multivariate Normal mixture was proposed by Day (1969), although maximum likelihood estimation of mixture density problems had been discussed in other work in the 1960s according to Redner and Walker (1984). This method was proposed as an improvement to the more difficult estimation by the methods of moments first introduced by Pearson (1894). An extensive evaluation of the method is found in Dempster, Laird, and Rubin (1977). Green (1984) discusses how this method compares to the method of iteratively reweighted least squares. Titterington, Smith, and Makov (1985) discuss finite mixture models in general along various techniques and procedures including EM.

Finite mixture models are relatively common in the economic literature. The switching regression model discussed by Goldfeld & Quandt (1972) might be considered a type of finite mixture model; although in this case, the resulting maximum likelihood

problem was solved by non-linear estimation methods other than EM. As was seen in the previous section, mixture models have also been used to explain the distribution of price changes by assuming the price sequence contains a subordinated process. Recent examples of the method include estimation of the joint distribution of prices and volatility (Danielsson (1994)), and partially adaptive regression models (Phillips (1994)).

Finite mixture models are also related to latent variable problems in that the variable identifying the mixture is often unobserved. Aitkin and Wilson (1980) review this approach to identify outliers in a mixed sample of "good and bad" observations in a sample. Their work serves as a simple introduction to the method and will demonstrate how the EM method will be used to identify a mix of "good and not so good" market prices.

Given a sample of independent observations, y_1 to y_n, it is assumed that a subset of these observations do not reflect the true population. The probability model consists of the true distribution and an alternate distribution along with a mixing proportion. The combined density function (from Aitkin and Wilson) is

$$f(y) = (p)(f_1(y)) + (1 - p)(f_2(y))$$

where $f_1()$ and $f_2()$ can be any density function. Normal density functions with equal variances and unique means are used for this example.

$$f_i = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-(y - \mu_i)^2 / 2\sigma^2 \right)$$

The parameters of the model are estimated by maximum likelihood and yield solutions,

$$\hat{p} = \sum_i \hat{P}(1 \mid y_i) / n$$

$$\hat{\mu}_j = \sum_i y_i \hat{P}(j \mid y_i) \Big/ \sum_i \hat{P}(j \mid y_i) \qquad j = 1, 2$$

$$\hat{\sigma}^2 = \sum_j \sum_i \left\{ (y_i - \hat{\mu}_j)^2 \hat{P}(j \mid y_i) \right\} \Big/ \; n \qquad j = 1, 2$$

where P() is the probability estimate for an observation to belong to one or the other subgroups. This estimate derives from the likelihood ratio of the subgroups and is defined as

$$\hat{P}(1 \mid y_i) = \hat{p}\ \hat{f}_1(y_i) \,/\, \left\{ \hat{p}\ \hat{f}_1(y_i) + (1 - \hat{p})\ \hat{f}_2(y_i) \right\}$$

The estimation of the model proceeds in similar way to reiteratively reweighted least squares. Initial estimates of P() are chosen, then the loglikelihood equation is maximized for each parameter.[2] These parameter estimates are used to adjust P() and the loglikelihood equation is again maximized. The algorithm continues, alternating between the estimation of the probability for each observation and the maximization of the loglikelihood equation, until convergence is reached. The criteria for convergence may be either that the value of the likelihood equation ceases to improve for some tolerance or that the parameter estimates stabilize for some predetermined tolerance.

Darwin's observations of heights of pairs of plants (*Zea Mays*) which were either self-pollinated or cross-pollinated has been used as sample data for techniques to detect outliers by Box and Tiao (1968) and Abraham and Box (1978). Aitkin and Wilson (1980) demonstrate how the EM algorithm can be used classify observations into two distributions: a true distribution and an outlier distribution. The original Darwin data discussed in Fisher (1935) is reproduced in Appendix A. A SAS/IML implementation of the EM algorithm is shown in Appendix B along with the results for the Darwin data. The

[2] According to a discussion in Everitt & Hand (1981), the estimation is not extremely sensitive to the initial estimates of the parameters. Also, it has been shown that for a univariate function, the likelihood value increases monotonically over iterations of the EM algorithm.

supposed outliers in the data are -67 and -48 which represent in eights of inches the difference between crossed and self-pollinated plant heights. All of the remaining differences in plant height are positive.

The results after 6 iterations are also reported in Appendix B. These results match the results found in Aitkin and Wilson. The final loglikelihood value is 143 and the mean of the outlier distribution is -57 whereas the mean of the supposed true distribution is 33. The probability that each of the negative observations may be classified as belonging to the outlier distribution is greater than .99 while the probability that any other observation may be classified as an outlier is less than .01.

3.2 Asset market simulation

Before applying the EM technique to experimental data, the technique is applied to simulated data to demonstrate how the algorithm performs under ideal conditions. The experimental Irmkt sessions used two Normal distributions with same means and unique variances (see Chapter 2). For this simulation, two distributions will be used where the mean of each distribution is $2.50, while the standard deviation is 10¢ for one distribution, and 50¢ for the second. These values are comparable to the moments used as signals in the experimental sessions. The two distributions are then mixed and the EM algorithm is applied to estimate the moments of each distribution. The identification of each observation predicted by the estimation is then compared with the true identity of each observation. This simulation might be described as allowing one trader to observe the signals for all of the traders, then estimating the mean value of the signals and the corresponding precisions.

The results of the simulation are shown in Appendix C. Each sub sample of the mix was comprised of 30 draws.[3] The estimated proportion of each signal is .42 vs. the actual value of .50. The mean of each group of signals is 2.485 and 2.518 vs. the actual values of 2.485 and 2.532. The estimated precision for each group is .109 and .408 vs. the

[3] Although the draws were made from a Normal distributions defined by identical means of 2.50 and standard deviations of .50 or .10, the actual sample moments were used for the comparison.

actual values of .109 and .442. The estimated classifications of the observations is also shown. The observations were sorted after the estimation: the first 30 observations belong to the first group, and the remaining 30 belong to the second group. Many of the predictions for the first group have probabilities of greater than .75 indicating that there is a high probability that these observations belong to the first group. The mean absolute deviation of the predicted probabilities vs. actual group classification is .429, indicating moderate predictive ability.

3.3 Information Criteria

Prediction of individual market orders is one measure of the usefulness of this model. Another criterion is how well the model explains the distribution of observed market orders. An increase in the information about the classification of market orders presumably allows a trader to better understand the underlying price process. Information in maximum likelihood estimation problems is usually measured by the Fisher information matrix (see, e.g., Green (1993)). Using the notation of Titterington, Smith, and Makov (1985), the Fisher information matrix for n observations is

$$n\, I(\psi) \;=\; E\left[D_\psi \mathcal{L}(\psi)\, D_\psi \mathcal{L}(\psi)^{\mathrm{T}} \right],$$

where ψ is a vector of parameters, $\mathcal{L}(\psi)$ is the loglikelihood equation, and D_ψ is the first derivative with respect to the parameters in the vector ψ. Since the expected value of this expression is difficult to calculate, an alternative estimator of the information matrix is often used. This is

$$n\, I(\hat{\psi}) \;=\; -E\left[D_\psi^2 \mathcal{L}(\hat{\psi}) \right],$$

where $\hat{\psi}$, the estimates of the loglikelihood equation at its maximum, replace the expected actual values. It is a well known property of the maximum likelihood technique that the inverse of the information matrix provides the asymptotic covariance matrix for the estimates. The details of this calculation of this matrix for the Darwin data are shown in Appendix B[4].

The information matrix may be interpreted as how much of the available information from the observations are captured in the model specification. This matrix may also be compared to a more general specification of the model. Specification testing in maximum likelihood problems is commonly performed with a loglikelihood ratio test where the maximized loglikelihood values from both specifications are compared. Aitkin & Wilson (1980) remark, however, that for finite mixture problems, the loglikelihood ratio test is not valid due to the non-regularity of the model. In common with many latent variable models, the mixing proportion is not identifiable in the general (non-mixed) specification.

While the ratio test is not available, the information matrix is still useful. Titterington, Smith, & Makov (1985) compare the information matrix associated with a fully categorized set of observations with information matrix associated with a mixture of observations. When assuming the fully categorized observations are more informative than the uncategorized observations, the gain in amount of information can be measured. This type of analysis might be thought of as a comparison of the entropy of two systems. In the case considered, a single Normal distribution would exhibit a higher quantity of entropy than an ordered (categorized) system. The relation Tittington, Smith, & Makov use is

$$I_c = I_u + I_e$$

where I_c is the information matrix from a fully categorized model, I_u is the information matrix from an uncategorized model, and I_e is the information matrix representing the

[4] As the analytical second derivatives are complex, and facilities for computing derivatives are absent in SAS, the derivatives were computed in Mathamatica and later transferred to the SAS program.

extra information associated from knowing the correct categorization of the observations. Since both I_c and I_u are positive semidefinite matrices when the loglikelihood function is evaluated at its maximum value, the difference of these matrices, I_e will also be a positive semidefinite matrix. Additional details are found in Titterington, Smith, & Makov (1985).

Viewing the estimation problem in terms of increasing the amount of information gathered from a set of observations gives insight into the equilibrium price equation derived in the rational expectations model of Chapter One. The equilibrium price equation was derived by assuming individual traders maximized a negative exponential utility function subject to a budget constraint to arrive at the individual demand for each trader. These individual demands were then summed over all traders and a market clearing condition was imposed so that demand equaled supply. The result was

$$p_t = \frac{\rho_o \psi_o + \mu \rho_t^{s1} \overline{y}_t^{\,1} + (1-\mu)\rho_t^{s2} \overline{y}_t^{\,2}}{\rho_o + \mu \rho_t^{s1} + (1-\mu)\rho_t^{s2}},$$

where p_t is the equilibrium price in period t, the mean and precision of the information common to both types of traders is ψ_o and ρ_o, the proportion of informed traders is μ, the mean and precision of the signals of the informed and uninformed traders are $\overline{y}_t^{\,1}$, ρ_t^{s1} and $\overline{y}_t^{\,2}$, ρ_t^{s2}.

The equivalence of the solution of the rational expectations problem and the proposed mixing models described here is a key insight. This equation determining the equilibrium price in the rational expectations model can be shown to also be the best linear unbiased (BLU) estimator of the mean value of a three component mixing model. The properties of this estimator are discussed in Bement & Williams (1969 p. 1375). The important quality of the BLU estimator is that it maximizes the available information from a set of observations. It can also be shown that the information derived from the observations using this estimator will be greater than the information associated with estimators from a

simpler (non-categorized) model. This increase in information follows the relation describing the extra information (I_e) describing the difference between a categorized model vs. an uncategorized model.

Extending the work of Williams (1967), Bement & Williams (1969) also derive an approximation for the variance of the weighted mean of a two component mixing model. They compare the finite series expansions of the estimator when sample data is used with the finite series expansion of the estimator when the variance of each component distribution is known. In a sense, this method of approximating the variance of the estimator measures the information lost when the information describing the categorization of observations (the variance of each category) is not available.

Bement & Williams use the variance approximation to compare the weighted average estimator to other possible estimators to describe the mean of the mix of observations. They develop criteria for selecting the optimal estimator based on the variance of the weighted average estimator. The alternate estimators suggested by Bement & Williams are the average sample mean, the pooled sample mean, and the mean with the smaller variance. These are

i. $\left(\overline{y}^{\,1} + \overline{y}^{\,2}\right) / 2$

ii. $\left(n_1 \overline{y}^{\,1} + n_2 \overline{y}^{\,2}\right) / \left(n_1 + n_2\right)$

iii. $\overline{y}^{\,1} \left(\sigma_1 / n_1 \leq \sigma_2 / n_2\right)$

where σ_1 and σ_2 are the variances of the two sub samples. It is proven that the weighted average variance approximation will be a superior estimator compared with these alternatives when the following inequalities are satisfied

i. $4Q < \left(1 / n_1 + 1 / n_2\right)$

ii. $Q < [n_1 + (n_2 / \rho)] / (n_1 + n_2)^2$

iii. $Q < 1 / n_1$

where $\rho = \sigma_1 / \sigma_2$, and Q is defined by

$$\frac{1}{n_2\rho + n_1} \times \left[1 + \frac{n_1 n_2 \rho}{(n_2\rho + n_1)^2}\right] \times$$

$$\left[\begin{array}{c} 2\left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}\right) \\ -\frac{16}{n_2\rho + n_1}\left[\frac{n_2\rho}{(n_1 - 1)^2} + \frac{n_1}{(n_2 - 1)^2}\right] + \\ \frac{12}{(n_2\rho + n_1)^2}\left[\frac{3n_2^2\rho^2}{(n_1 - 1)^2} + \frac{n_2^2\rho^2 - 4n_1 n_2 \rho + n_1^2}{(n_1 - 1)(n_2 - 1)} + \frac{3n_1^2}{(n_2 - 1)^2}\right] \\ + \left[\frac{12\, n_2^2\rho^2}{(n_1 - 1)^3} + \frac{12\, n_1^2}{(n_2 - 1)^3}\right] \end{array}\right].$$

The approximation to the variance of weighted mean estimator is then defined by adjusting the variance of the first sub sample by the factor defined above. This is

$$\sigma_w = Q \times \sigma_1,$$

where σ_w is the variance of the mean of the mix.

The variance approximation is useful in terms of a trading model because it allows

the optimal estimator of the mean of a mix of signals to be identified, and using this optimal estimator gives the best estimate of the equilibrium price defined in Chapter One. The mixing model therefore provides two kinds of information: the first is classification of individual observations according to trader type, and the second is an estimate of the equilibrium price function. The simulations showed that the EM algorithm performs reasonably well in predicting the moments of the distributions, and in identifying individual observations. In the next section, these techniques are applied to actual experimental market data from Chapter Two.

4 Application to Experimental Data

4.1 Description of the model

The following model employs the estimation maximum likelihood (EM) algorithm to estimate a mix of two types of signals simply by observing market orders. Each trader receives only a single private signal at the beginning of a trading period, and each additional price observation (bid or ask) within the period is considered an additional signal of unknown precision. The precision of these additional signals is in fact a mix of two precisions: the precisions of informed and uninformed trader. The model identifies the moments of each of the distributions in this mix, and identifies the distribution from which each price observation is drawn. As a result, the model demonstrates how larger volumes of market orders improves the estimation of the precision of signals of the informed traders, and thereby improves the estimation of equilibrium prices. Volume is related to the estimation the true asset worth in that it measures the number of price observations and determines the sample size for the estimation.

It is assumed that trader behavior can be modeled as ZI in that traders add a random profit to their signal to determine a bid value, and subtract a random profit to their signal to determine an ask value. Also, it is assumed that the profit margins can be considered to be

Normally distributed with a positive mean value, and this mean value is the same across traders. Finally, it is assumed that the sequence of market orders is random, and gives no indication of the type of trader submitting the order.

The observed price sequence of bids and asks in levels given these assumptions is therefore modeled as a mix of two distributions and a stable bid-ask spread. The profit margin for either side of the market will shift the distribution of signals outward away from the mean value from which the signals are drawn, although the mean value of the both distributions will be the same. The resulting sequence of bids levels will then be a mixed distribution of Normals with a common mean, and the same is true of the sequence of ask levels. The mean value of the bid sequence and the ask sequence will not be equal, although the difference in means will be Normally distributed.

The five parameters to be estimated are the mixing proportion for the two groups (p), the means of each groups (μ_1, μ_2), and the variances for each group (σ_1, σ_2). The loglikelihood function is

$$\log li = \sum_i [\ln [(p) f_1 (\mu_1, \sigma_1) + (1 - p) f_2 (\mu_2, \sigma_2)]] .$$

Once the moments of the two distributions are identified, these distributions are assigned to either the informed or uninformed group of traders. In order to assign the distributions it will be assumed that informed traders typically outbid (outask) uninformed traders. This is equivalent to assuming informed traders are more likely to have the inside market (the highest bid or lowest ask). Since traders must have the inside market to complete a trade, and profits can only be earned through transactions, it is to the advantage of a trader to have the inside market. It is assumed that the information advantage of the insiders allows them to capture the inside market. Using this assumption, the distribution with the greater (lesser) mean will be assigned to the bid (ask) of the informed traders. These assumptions are tested below.

4.2 Testable Hypothesis

The mixture of distributions (MOD) model is tested on laboratory data according to two criteria: the identification of market orders and the estimation of equilibrium price. The null hypothesis is that market orders (either bids or asks) are equally likely from informed and uninformed traders, and a strategic trader is unable to identify the type of trader using market orders. The alternate hypothesis is that the MOD model allows traders to identify the type of trader submitting market orders, and provides an estimation of the equilibrium price.

Since there are only two types of traders considered, the model is only useful if it outperforms predictions by a random variable drawn from a binomial distribution. Define the state as a binary variable where Informed = 1, and uninformed = 0. The expected value of a single draw from a binomial distribution is simply the probability of a success where success is defined as choosing the correct state. The expected error is the actual state less the expected value. Given that the state takes only two values, the expected error is always .5. To test the mixture of distributions hypothesis, the predictions of the model will be compared against this benchmark.

It should be noted that rejection of the MOD model does not necessarily imply rejection of the ZI assumption. Traders may still be using a ZI strategy to place bids and offers while the resulting market orders are so similar that the MOD model cannot correctly identify them.

4.3 Results

4.3.1 Test of Normality

The experimental data exhibits market orders which are not Normally distributed. The Shapiro-Wilk test of Normality is applied to the observed bids and asks from sessions

9 through 17 (Table 1).[5] Each side of the market is analyzed independently. On the bid side, the Normality test for the combined informed/uninformed group (group 9) could be rejected about 77% of the time indicating that the bids from the mix of traders could not be considered to be Normal. For the informed biders, the number of times Normality can be rejected drops to 61% of the time, and for the uninformed to 60% of the time. For the ask side, the results are comparable. Normality for the asks of the combined group of informed/uninformed traders is rejected more often than for the groups considered independently.

4.3.2 Characterization of Bids and Asks

In Table 2, descriptive statistics are run on several time series of prices (bid, logbid, bought, ask, logask sold). Several features of these time series are apparent. Informed traders tend to bid higher than uninformed traders and ask below uninformed traders. This is seen in the mean bid within each period of each session. A paired t-test was run for all the periods. The null hypothesis of same means across groups can be rejected at the .01 level for the bid or the ask side.[6] Session 15 is shown as a sample of the complete data. The same is true if the log of bids or asks is taken.

The variance of the bids or asks across groups may or may not be the same across the two groups. A difference in the mean variance was computed for a simple comparison, then an F-test of the variances was performed taking into account the degrees of freedom for each group in each period. These results are presented for a sample session (Session 15). A summary of the significance of this test is also shown. For all periods in all sessions, the null hypothesis of same variances across groups can be rejected at the 10% level in about 38% of the periods examined.

For the actual transaction prices (bought or sold), there appears to be little difference

[5] Conover (1980) discusses the theory underlying the Shapiro-Wilk test along with alternative Normality tests. The Univariate Proc in SAS was used to perform the tests reported here.

[6] The number of observations is calculated as the number of sessions times the number of periods in each session.

between the two groups in either the mean values or variances. For this reason, transaction prices were excluded from the data and only market orders (unaccepted bids and asks) were used.

4.4 Estimates of the model

Exploiting the differences in the observed behavior of the informed vs. uninformed traders, a five parameter model (p, mu1, mu2, sigma1, sigma2) is compared with a simple two parameter model (mu, sigma). It is then shown that by assigning the means in this model to the informed and uninformed groups, a strategic trader could do better using this model than by averaging of all observed prices. The buy side and the sell side are modeled independently.

Session 16 period 1 is used as an example. The actual mean bid of Group 1 (informed traders) is 1.76 vs 2.00 for Group 2 (uninformed traders). The model assumes two Normal distributions with unique means and variances. As shown in Table 3a, convergence took place after 15 iterations (k=15). The value of the likelihood function is given with and without a constant (log(-2*Pi)), as well as the estimated parameters of the model. The estimated mean for group 1 of 1.67 compares with the actual mean of 1.76, and the estimated mean for group 2 of 2.11 compares with the actual mean of 2.00. The variances are also comparable.

Since probabilities for each observation are given, these are compared with the actual group classification of each observation. The vector **s** defines the predicted probabilities for each observation. The error for each observation is computed and the mean for all 40 observations is given. Since each observation could be considered a binomial draw from either distribution, the mean error would be .50 if it were equally likely that any observation belonged to either group. In the model, the mean absolute error is .29. It is unlikely that this value could be produced by randomly assigning observations to groups.

The mixing model is estimated for all periods of all sessions in Table 4. For most of the sessions, the mean absolute deviation is less than .50, indicating that the model does

slightly better than a binomial draw. This improvement allows the null hypothesis of no improvement to be rejected, and the alternate MOD hypothesis to be accepted. Results after the data were filtered are also reported. Since the mixing model must discriminate between two groups of data, the model will perform best when these data are distinctly different while the model will have difficulty when these data are very similar. A filter is used to eliminate the periods where the predicted difference between the data is less than 20¢. As seen in Table 4 there is an improvement in the predictions of the mixing model when this filter is used.

4.5 Estimation of Equilibrium Price

The estimates of the mixing model can now be compared to the expected equilibrium price. The experimental sessions provided one signal per trader at the beginning of each period, and traders were split equally between the informed and uninformed groups. One method of computing the expected equilibrium price is to fully aggregate information by taking into account each trader's signal. This method was used in Chapter Two. The actual market orders, however, were voluntary so some traders were over represented. A second method considers only the signals of the active traders. To calculate the average signal for the informed trader, each market order by an informed trader contributes one observation. The same rule is used to calculate the average signal for the uninformed traders. And lastly, while the traders were initially assigned equally between the two trader types, the actual participation rate is used for the proportion of informed and uninformed traders. The expected equilibrium price is then calculated according the the equilibrium price function using the information from one of the above aggregation methods. The resulting expected equilibrium price can then be used as a benchmark to test the mixing model.

The estimates of the mixing model provide an estimate of the mean for both types of traders along with the mixing proportion. The variance approximation is also calculated and the weighted average mean of the mixing model can be compared to alternative

estimators of the mean of the mix of observations using the criteria discussed in section 3.3. An additional estimator was also considered. Since it was found that informed traders tend to bid higher (ask lower) than the uninformed traders, the largest (smallest) of the two means was considered as an estimator.

The results of the comparison are shown in Table 5 & 6 for each experimental session. The actual trader signals weighted by the actual signal sample variances are used as the baseline for the comparison in Table 5, and the aggregated information is used for the comparison in Table 6. The mean absolute deviation (MAD) of the observed market price and of each estimator with respect to this benchmark is shown. The MAD for the optimal estimator using the Q criteria discussed in section 3.3 is also shown. For many of the sessions, the optimal estimator shows a smaller MAD than the observed market price. The weighted average mean is often the best estimator, and the Q criteria indicates many cases when alternate estimators are optimal. The largest (smallest) of the means for bids (asks) used as an estimator improves upon all other estimators in many of the sessions. These results support the mixing of distributions hypothesis in that the estimated of the model provide more information than simply observing price.

5 Discussion

The model of a mixture of Normal distributions presented here allows a role for volume in each trader's estimation of the current fundamental. Volume increases the number of sample observations in a maximum likelihood estimation, and may improve a trader's estimation of equilibrium price. Unlike the model of Blume, Easley, & O'Hara (1994) volume here might directly enter into the demand function of a strategic trader.

This model is fairly simple, and knowledge of the structure of the market could be used to enhance the model. Transaction prices were not used although it is known that these valuable provide information. The bid side and the ask side are modeled independently even though the same signal allows each trader to be active on both sides of

the market. A more general model which includes both sides of the market in one estimation might improve the estimation. Another enhancement might consider the convergence over time of market prices to the true value. In our model, price observations early and late in the period are treated equally. Also it is known that the variance of price changes declines over time, taking this into account would allow a better estimation of the variance of the original signals.

As was seen in the introduction, this type of model has wide applicability. The resolution of the sources of price variability would be of great importance in all types of financial markets. Hopefully it has been demonstrated how key structural features of a market such as the observed behavior characteristics of two types of traders can be incorporated directly into a mixture of distributions model.

References

Bachelier, L., 1900, "Theory of speculation," reprinted in P. Cootner (ed.), 1964, *The Random Character of Stock Market Prices*, Cambridge: MIT Press. p. 17-78.

Barnett, Vic. and Toby Lewis, 1994, *Outliers in Statistical Data*, 3rd edition, New York: John Wiley & Sons.

Bement, T.R., and J.S. Williams, 1969, "Variance of weighted regression estimators when sampling errors are independent and heteroscedastic," *American Statistical Association Journal*, 64. p. 1369-1382.

Blume, Lawrence, David Easley, and Maureen O'Hara, 1994, "Market statistics and technical analysis: the role of volume," *Journal of Finance*, XLIX:1. p. 153-181.

Clark, Peter, 1973, "A subordinated stochastic process model with finite variance for speculative prices," *Econometric*, 41:1. p. 135-159.

Conrad, Jennifer, Allaudeen Hameed, and Cathy Niden, 1994, "Volume and autocovariances in short-horizon individual security returns," *Journal of Finance*, XLIX:4. pg. 1305-1329.

Conover, W.J., 1980, *Practical Nonparametric Statistics*, 2nd ed. New York: John Wiley & Sons.

Easley, D., and J.O. Ledyard, 1993, Theories of price formation and exchange in oral auctions," in *The Double Auction Market: Institutions, Theories, and Evidence*, edited by Friedman and Rust, New York: Addison-Wesley. p. 63-97.

Fisher, Ronald A., 1935, *The Design of Experiments*, 9th edition 1971 reprinted 1974, New York: Hafner Press.

Friedman, D., 1991, "A simple testable model of double auction markets," *Journal of Economic Behavior & Organization*, p. 47-70.

Friedman, Daniel and Stoddard Vandersteel, 1980, "Short-run fluctuations in foreign exchange rates: an exploration of the data," UCLA discussion paper number 171.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen, 1992, "Stock prices and volume," *Review of Financial Studies*, 5. p. 199-242.

Gode, Dhananjay, and Shyam Sunder, 1993, "Lower bounds for efficiency of surplus extraction in double auctions," in *The Double Auction Market: Institutions, Theories, and Evidence*, edited by Friedman and Rust, New York: Addison-Wesley. p. 199-219.

Green, William H., 1993, *Econometric Analysis*, New York: Macmillan Publishing Co.

Karpoff, Jonathan, 1986, "A theory of trading volume," *Journal of Finance*, XLI:5. p. 1069-1087.

Karpoff, Jonathan, 1987, "The relation between price changes and trading volume: a survey," *Journal of Financial and Quantitative Analysis*, 22. p. 109-126.

Park, Hun Y., 1993, "Trading mechanisms and price volatility: spot versus futures," *Review of Economics and Statistics*. p. 175-179.

Roa, C. Radhakrishna, 1965, *Linear Statistical Inference and Its Applications*, New York: John Wiley & Sons.

Tauchen, George and Mark Pitts, 1983, "The price variability-volume relationship on speculative markets," *Econometrica*, 51:2. p. 485-505.

Taylor, Stephen, 1985, *Modeling Financial Time Series*, New York: John Wiley & Sons. [see chapter 2, "Features of Financial Returns," especially p. 44 on kurtosis.]

Upton, David and Donald Shannon, 1979, "The stable paretian distribution, subordinated stochastic processes, and asymptotic lognormality: an empirical investigation," *Journal of Finance*, XXXIV:4. p 1031-1039.

Wang, Jiang, 1994, "A model of competitive stock trading volume," *Journal of Political Economy*, 102:1. p. 127-168.

Williams, J.S., 1967, "The variance of weighted regression estimators," *Journal of the American Statistical Association*, 62. p. 1290-1301.

Wilson, R. B., 1987, "On equilibria of bid-ask markets," in *Arrow and the Ascent of Modern Economic Theory*, edited by G. Feiwell, New York: MacMillan. p. 375-414.

References specific to maximum likelihood and the EM algorithm

Abraham, B. and G.E.P Box, 1978, "Linear models and spurious observations," *Applied Statistics*, 27. p. 131-138.

Aitkin, Murray and Granville Tunnicliffe Wilson, 1980, "Mixture models, outliers, and the EM algorithm," *Technometrics*, 22:3. pg. 325-331.

Box, G.E.P. and G.C. Tiao, 1968, "A Bayesian approach to some outlier problems," *Biometrika*, 55. p. 19-129.

Danielsson, Jon, 1994, "Stochastic volatility in asset prices: estimation with simulated maximum likelihood," *Journal of Econometrics*. 64. p. 375-400.

Day, N.E., 1969, "Estimating the components of a mixture of normal distributions," *Biometrica*, 56:3. p. 463-474.

Demster, A. P., N. M. Laird and D. B. Rubin, 1977, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, B*., 39.

p. 1-38.

Everitt, B.S. and D.J. Hand, 1981, *Finite Mixture Distributions*, New York: Chapman and Hall.

Goldfeld, Stephen and Richard E. Quandt, 1972, *Nonlinear Methods in Econometrics*, vol. 77 in *Contributions to Economic Analysis* edited by Jorgenson and Waelbroeck. Amsterdam: North-Holland.

Green, P.J., 1984, "Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives," *Journal of the Royal Statistical Society B*. 46. p. 149-192.

Pearson, K. 1894, "Contributions to the mathematical theory of evolution," *Phil. Trans. R. Soc*. 184. p. 71-110.

Phillips, Robert F., 1994, "Partially adaptive estimation via a normal mixture," *Journal of Econometrics*. 64. p. 123-144.

Redner, Richard A. and Homer F. Walker, 1984, "Mixture densities, maximum likelihood and the EM algorithm," *Siam Review*, 26:2. p. 195-239.

Titterington, E. M., A.F.M. Smith, and U.E. Makov, 1985, *Statistical Analysis of Finite Mixture Distributions*, New York: John Wiley & Sons.

Appendix A

Darwin Data for Heights of *Zea Mays*

Pot Number	Crossed	Self-Fert.	Difference
I.	23.5	17.375	6.125
	12	20.375	-8.375
	21	20	9
II.	22	20	8
	19.125	18.375	0.75
	21.5	18.625	2.875
III.	22.125	18.625	3.5
	20.375	15.25	5.125
	18.25	16.5	1.75
	21.625	18	3.625
	23.25	16.25	7
IV.	21	28	3
	22.125	12.75	9.375
	23	15.5	7.5
	12	18	-6

Notes: Data is reproduced from Fisher (1935). Differences in plant heights are converted to eights of an inch for the analysis to correspond with Aitkin & Wilson (1980).

Appendix B

EM Algorithm in SAS/IML Applied to Darwin Data

Iteration History

K	LF2	LF3	SPH	MU1	MU2	SGS
1	147.3486	119.78044	0.0666667	-67	27.214286	777.35714
2	144.54424	116.9761	0.1098303	-58.70257	30.75891	547.19352
3	143.57921	116.01105	0.1319683	-57.26193	32.821484	400.06512
4	143.5671	115.99896	0.1335204	-57.34806	32.996127	385.36988
5	143.5671	115.99895	0.1335125	-57.37018	32.998713	384.90068
6	143.5671	115.99895	0.1335109	-57.37141	32.998733	384.8842

Prediction Probability Vector

S						
0.9999833	0.9985557	0.0021495	0.001345	0.0003291	0.0002058	0.0000398
: 0.0000315	0.0000123	9.7255E-6	5.8111E-7	8.8815E-8	1.7167E-8	6.7112E-9
: 1.983E-10						

Standard Errors of the Estimated Parameters

SPH_SE	MU1_SE	MU2_SE	SGS_SE
0.0879648	14.029329	5.4484765	3.6317029

Covariance and Inverse Covariance Matrices

COV			
-129.2449	0.0063644	0.0049521	0.0065059
0.0063644	-0.005082	0.000052	0.0003169
0.0049521	0.000052	-0.033687	-0.00014
0.0065059	0.0003169	-0.00014	-0.07584

INVCOV			
0.0077378	0.0097452	0.0011496	0.0007024
0.0097452	196.82208	0.3020679	0.8227611
0.0011496	0.3020679	29.685896	-0.053506
0.0007024	0.8227611	-0.053506	13.189266

Appendix B

Program Listing for Darwin Data

```
  data main1 ;
  title 'fn:kov.sas - Detection of Outliers in Darwin Data - CGP 1995' ;
  input d @@ ;
  cards ;
  -67 -48 6 8 14 16 23 24 28 29 41 49 56 60 75
  ;

proc iml ;
  use main1 ;
  read all into y ;

  pi = 3.14159265 ;
  n  = ncol(y`) ;
  d  = ({1} || j(1,n-1,0))` ;
  i  = j(1,n,1) ;

  start f(y,mu,s) ;
  fv = (1/(s*sqrt(2*3.14159265)))*exp((-(y-mu)##2)/(2*s##2)) ;
  return(fv) ;
  finish ;

  s = d ;
  k = 0 ;
  lf2 = 0 ;

do until((abs(lf0-lf2)<.00001) | (k>10)) ;

  sph = (i*s)/n ;
  mu1 = (y`*s)/(i*s) ;
  mu2 = (y`*(1+(-1*s)))/(i*(1+(-1)*s)) ;
  sg  = sqrt(((((y-mu1)##2)`*s)+(((y-mu2)##2)`*(1+(-1)*s)))/n) ;

  s   = sph*f(y,mu1,sg) / (sph*f(y,mu1,sg)+(1-sph)*f(y,mu2,sg)) ;

  lf0 = lf2 ;
  lf1 = (sph*f(y,mu1,sg) + (1-sph)*f(y,mu2,sg)) ;
  lf2 = -2*(i*log(lf1)) ;
  lf3 = lf2 - (15*log(2*pi)) ;

  k = k + 1 ;

  sgs = sg**2 ;
  print k lf2 lf3 sph mu1 mu2 sgs ;

end ;

  s = s` ;
  print s ;

* add analytical derivatives ;
```

```
start fd11(y,sphes,mu1es,mu2es,sges,pi) ;
d11 = -((1/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) -
         1/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2/
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2) ;
return(d11) ;
finish ;


start fd44(y,sphes,mu1es,mu2es,sges,pi) ;
d44 = -  ((((y - mu2es)##2*(1 - sphes))/
          (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
         (1 - sphes)/
          (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2) +
         ((y - mu1es)##2*sphes)/
          (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2))##2/
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2) +
   (((y - mu2es)##4*(1 - sphes))/
       (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##7) -
      (5*(y - mu2es)##2*(1 - sphes))/
       (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##5) +
      ((2/pi)##(1/2)*(1 - sphes))/(exp((y - mu2es)##2/(2*sges##2))*sges##3) +
      ((y - mu1es)##4*sphes)/
       (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##7) -
      (5*(y - mu1es)##2*sphes)/
       (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##5) +
      ((2/pi)##(1/2)*sphes)/(exp((y - mu1es)##2/(2*sges##2))*sges##3))/
    ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
      sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges)) ;
return(d44) ;
finish ;


start fd22(y,sphes,mu1es,mu2es,sges,pi) ;
d22 = -((y - mu1es)##2*sphes##2)/
    (2*exp((y - mu1es)##2/sges##2)*pi*sges##6*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2) +
   ((y - mu1es)##2*sphes)/
    (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##5*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) -
   sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) ;
return(d22) ;
finish ;


start fd33(y,sphes,mu1es,mu2es,sges,pi) ;
d33 = -((y - mu2es)##2*(1 - sphes)##2)/
    (2*exp((y - mu2es)##2/sges##2)*pi*sges##6*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2) +
```

```
   ((y - mu2es)##2*(1 - sphes))/
    (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##5*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) -
   (1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) ;
return(d33) ;
finish ;


start fd12(y,sphes,mu1es,mu2es,sges,pi) ;
d12 = -(((y - mu1es)*(1/
           (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) -
          1/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))*sphes)/
      (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
        ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
           sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2)) +
   (y - mu1es)/
    (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) ;
return(d12) ;
finish ;


start fd13(y,sphes,mu1es,mu2es,sges,pi) ;
d13 = -(((y - mu2es)*(1/
           (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) -
          1/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))*(1 - sphes))/
      (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
        ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
           sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2)) -
   (y - mu2es)/
    (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) ;
return(d13) ;
finish ;


start fd14(y,sphes,mu1es,mu2es,sges,pi) ;
d14 = -(((1/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) -
          1/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))*
        (((y - mu2es)##2*(1 - sphes))/
           (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
          (1 - sphes)/
           (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2) +
          ((y - mu1es)##2*sphes)/
           (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
          sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2)))/
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2) +
   ((y - mu1es)##2/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
      (y - mu2es)##2/
       (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
```

```
      1/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2) +
      1/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2))/
    ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
      sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges)) ;
return(d14) ;
finish ;


start fd24(y,sphes,mu1es,mu2es,sges,pi) ;
d24 = -(((y - mu1es)*sphes*
        (((y - mu2es)##2*(1 - sphes))/
           (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
          (1 - sphes)/
           (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2) +
          ((y - mu1es)##2*sphes)/
           (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
          sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2)))/
      (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
        ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
           sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2)) +
   ((y - mu1es)##3*sphes)/
    (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##6*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) -
   (3*(y - mu1es)*sphes)/
    (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) ;
return(d24) ;
finish ;


start fd34(y,sphes,mu1es,mu2es,sges,pi) ;
d34 = -(((y - mu2es)*(1 - sphes)*
        (((y - mu2es)##2*(1 - sphes))/
           (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
          (1 - sphes)/
           (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2) +
          ((y - mu1es)##2*sphes)/
           (exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4) -
          sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##2)))/
      (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##3*
        ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
           sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2)) +
   ((y - mu2es)##3*(1 - sphes))/
    (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##6*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) -
   (3*(y - mu2es)*(1 - sphes))/
    (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges##4*
      ((1 - sphes)/(exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
        sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))) ;
return(d34) ;
finish ;

start fd23(y,sphes,mu1es,mu2es,sges,pi) ;
```

```
d23 = -(exp(-(y - mu1es)##2/(2*sges##2) - (y - mu2es)##2/(2*sges##2))*
       (y - mu1es)*(y - mu2es)*(1 - sphes)*sphes)/
    (2*pi*sges##6*((1 - sphes)/
          (exp((y - mu2es)##2/(2*sges##2))*(2*pi)##(1/2)*sges) +
         sphes/(exp((y - mu1es)##2/(2*sges##2))*(2*pi)##(1/2)*sges))##2) ;
return(d23) ;
finish ;

sumd11 = 0 ; sumd12 = 0 ; sumd13 = 0 ; sumd14 = 0 ;
sumd21 = 0 ; sumd22 = 0 ; sumd23 = 0 ; sumd24 = 0 ;
sumd31 = 0 ; sumd32 = 0 ; sumd33 = 0 ; sumd34 = 0 ;
sumd41 = 0 ; sumd42 = 0 ; sumd43 = 0 ; sumd44 = 0 ;

do m = 1 to n ;

  value = y[m] ;

  sumd11 = sumd11 + fd11(y[m],sph,mu1,mu2,sg,pi) ;
  sumd12 = sumd12 + fd12(y[m],sph,mu1,mu2,sg,pi) ;
  sumd13 = sumd13 + fd13(y[m],sph,mu1,mu2,sg,pi) ;
  sumd14 = sumd14 + fd14(y[m],sph,mu1,mu2,sg,pi) ;

  sumd22 = sumd22 + fd22(y[m],sph,mu1,mu2,sg,pi) ;
  sumd23 = sumd23 + fd23(y[m],sph,mu1,mu2,sg,pi) ;
  sumd24 = sumd24 + fd24(y[m],sph,mu1,mu2,sg,pi) ;

  sumd33 = sumd33 + fd33(y[m],sph,mu1,mu2,sg,pi) ;
  sumd34 = sumd34 + fd34(y[m],sph,mu1,mu2,sg,pi) ;

  sumd44 = sumd44 + fd44(y[m],sph,mu1,mu2,sg,pi) ;
end ;

* initiate matrix then assign values.  note symmetry ;

cov = I(4) ;

cov[1,1]=sumd11 ; cov[1,2]=sumd12 ; cov[1,3]=sumd13 ; cov[1,4]=sumd14 ;
cov[2,1]=sumd12 ; cov[2,2]=sumd22 ; cov[2,3]=sumd23 ; cov[2,4]=sumd24 ;
cov[3,1]=sumd13 ; cov[3,2]=sumd23 ; cov[3,3]=sumd33 ; cov[3,4]=sumd34 ;
cov[4,1]=sumd14 ; cov[4,2]=sumd24 ; cov[4,3]=sumd34 ; cov[4,4]=sumd44 ;

invcov = inv(-cov) ;

sph_se = sqrt(invcov[1,1]) ;
mu1_se = sqrt(invcov[2,2]) ;
mu2_se = sqrt(invcov[3,3]) ;
sgs_se = sqrt(invcov[4,4]) ;

print sph_se mu1_se mu2_se sgs_se ;
print cov ;
print invcov ;
```

Appendix C

Simulation Results

Input Data

Nobs	Variable	N	NMISS	MEAN	STD
60	Y1	30	30	2.48477	0.10891
	Y2	30	30	2.52324	0.44245
	Y	60	0	2.50401	0.32004

Iteration History

OBS	N	K	LF2	LF3	SPH	MU1	MU2	SG1	SG2
1	60	1	22.921	-4.647	0.500	2.485	2.523	0.107	0.435
2	60	2	22.582	-4.987	0.478	2.485	2.521	0.111	0.425
3	60	3	22.498	-5.070	0.466	2.486	2.520	0.113	0.421
4	60	4	22.468	-5.101	0.458	2.486	2.519	0.113	0.418
5	60	5	22.452	-5.116	0.452	2.486	2.519	0.113	0.416
6	60	6	22.442	-5.126	0.448	2.486	2.519	0.112	0.414
7	60	7	22.435	-5.133	0.445	2.486	2.518	0.112	0.413
8	60	8	22.430	-5.138	0.442	2.486	2.518	0.112	0.413
9	60	9	22.426	-5.142	0.440	2.486	2.518	0.111	0.412
10	60	10	22.423	-5.145	0.438	2.486	2.518	0.111	0.411
11	60	11	22.421	-5.147	0.436	2.486	2.518	0.111	0.411
12	60	12	22.419	-5.149	0.435	2.486	2.518	0.110	0.410
13	60	13	22.418	-5.150	0.433	2.486	2.518	0.110	0.410
14	60	14	22.417	-5.151	0.432	2.486	2.518	0.110	0.410
15	60	15	22.417	-5.152	0.431	2.486	2.518	0.110	0.409
16	60	16	22.416	-5.152	0.431	2.486	2.518	0.110	0.409
17	60	17	22.416	-5.153	0.430	2.486	2.518	0.109	0.409
18	60	18	22.415	-5.153	0.429	2.486	2.518	0.109	0.409
19	60	19	22.415	-5.153	0.429	2.485	2.518	0.109	0.409
20	60	20	22.415	-5.153	0.428	2.485	2.518	0.109	0.408
21	60	21	22.415	-5.153	0.428	2.485	2.518	0.109	0.408
22	60	22	22.415	-5.153	0.428	2.485	2.518	0.109	0.408

Predicted Probabilities (G=Group, Y=Observation, S=Prediction, E=Error)

OBS	INDEX	G	Y	S	E
1	11	1	2.30652	0.70704	0.29296
2	12	1	2.31994	0.67642	0.32358
3	13	1	2.34472	0.71800	0.28200
4	14	1	2.34969	0.57081	0.42919
5	15	1	2.35195	0.71647	0.28353
6	16	1	2.40306	0.73686	0.26314
7	17	1	2.40521	0.20611	0.79389
8	18	1	2.41964	0.72067	0.27933
9	19	1	2.42080	0.49133	0.50867
10	20	1	2.42389	0.71978	0.28022
11	22	1	2.42675	0.71006	0.28994
12	24	1	2.43058	0.66377	0.33623
13	25	1	2.43302	0.71270	0.28730
14	26	1	2.44992	0.73650	0.26350
15	27	1	2.47261	0.68622	0.31378
16	29	1	2.48225	0.29508	0.70492
17	30	1	2.49065	0.72714	0.27286
18	31	1	2.50613	0.71285	0.28715
19	32	1	2.51939	0.73723	0.26277
20	33	1	2.52940	0.60109	0.39891
21	34	1	2.53060	0.00090	0.99910
22	35	1	2.53481	0.25495	0.74505
23	36	1	2.54859	0.48733	0.51267
24	37	1	2.56980	0.38740	0.61260
25	38	1	2.57786	0.03204	0.96796
26	39	1	2.57990	0.05485	0.94515
27	43	1	2.63380	0.64171	0.35829
28	44	1	2.64818	0.00001	0.99999
29	48	1	2.70363	0.00027	0.99973
30	49	1	2.72978	0.23857	0.76143
31	1	2	1.50669	0.49836	0.49836
32	2	2	1.58658	0.58934	0.58934
33	3	2	1.90177	0.53568	0.53568
34	4	2	1.97146	0.70586	0.70586
35	5	2	1.97397	0.73323	0.73323
36	6	2	2.16792	0.73731	0.73731
37	7	2	2.18882	0.45399	0.45399
38	8	2	2.24610	0.66034	0.66034
39	9	2	2.25091	0.70344	0.70344
40	10	2	2.28768	0.71602	0.71602
41	21	2	2.42577	0.68907	0.68907
42	23	2	2.42691	0.58368	0.58368
43	28	2	2.47978	0.72896	0.72896

44	40	2	2.58444	0.00071	0.00071
45	41	2	2.59018	0.00000	0.00000
46	42	2	2.60913	0.00000	0.00000
47	45	2	2.64941	0.00227	0.00227
48	46	2	2.67540	0.16853	0.16853
49	47	2	2.69384	0.00000	0.00000
50	50	2	2.74229	0.39769	0.39769
51	51	2	2.82004	0.33055	0.33055
52	52	2	2.88731	0.08714	0.08714
53	53	2	2.90976	0.00001	0.00001
54	54	2	2.93667	0.00011	0.00011
55	55	2	2.93674	0.00469	0.00469
56	56	2	2.94314	0.65241	0.65241
57	57	2	2.96938	0.00013	0.00013
58	58	2	2.98811	0.00090	0.00090
59	59	2	3.05585	0.71182	0.71182
60	60	2	3.29127	0.00010	0.00010

Prediction Probabilities Mean Absolute Deviation

Mean
0.4291364

Appendix C

Simulation Program Listing

```
proc iml ;

* define signals for informed and uninformed traders ;

  seed = 131406 ;
  size = 30 ;

  y1 = 2.5 + (.10)*normal(repeat(seed,1,size)) ;
  g1 = repeat(1,1,size) ;
  y2 = 2.5 + (.50)*normal(repeat(seed,1,size)) ;
  g2 = repeat(2,1,size) ;

  y  = (y1 || y2)` ;
  g  = (g1 || g2)` ;

  create main1 var{y1,y2,y,g} ;
  append ;
  close main1 ;

* do stats on actual values ;

  use main1 ;
  summary var{y1 y2 y} stat{n nmiss mean std} ;
  close main1 ;

* shuffle observations ;

  sort main1 by y ;

  n = ncol(y`) ;
  d = ( j(1,(n-int(n/2)),1) || j(1,int(n/2),0) )` ;
  i = j(1,n,1) ;

* define normal distribution function ;

  start f(y,mu,s) ;
  fv = (1/(s*sqrt(2*3.14159)))*exp((-(y-mu)##2)/(2*s##2)) ;
  return(fv) ;
  finish ;

  s = d ;
  k = 0 ;
  lf2 = 0 ;

create main2 var{n,k,lf2,lf3,sph,mu1,mu2,sg1,sg2} ;

do until((abs(lf0-lf2)<.0001) | (k > 60) ) ;

  sph = (i*s)/n ;
  mu1 = (y`*s)/(i*s) ;
```

```
  mu2 = (y`*(1+(-1*s)))/(i*(1+(-1)*s)) ;
  sg1 = sqrt(((((y-mu1)##2)`*s) )/(n*sph)) ;
  sg2 = sqrt(((((y-mu2)##2)`*(1+(-1*s))) )/(n*(1-sph))) ;

  s   = sph*f(y,mu1,sg1) / (sph*f(y,mu1,sg1)+(1-sph)*f(y,mu2,sg2)) ;

  lf0 = lf2 ;
  lf1 = (sph*f(y,mu1,sg1) + (1-sph)*f(y,mu2,sg2)) ;
  lf2 = -2*(i*log(lf1)) ;
  lf3 = lf2 - (15*log(2*3.14159)) ;

  k = k + 1 ;

append var{n,k,lf2,lf3,sph,mu1,mu2,sg1,sg2} ;

end ;

close main2 ;

create main3 ;
append var{s} ;
close main3 ;

data main4 ;
  set main1 ;
  index = _N_ ;
  keep index g y ;
data main5 ;
  set main3 ;
  index = _N_ ;
  keep index s ;

data main6 ;
  merge main4 main5 ;
  by index ;
  if g = 1 then e = 1-s ;
  if g = 2 then e =   s ;

proc sort ;
  by g ;

proc print data=main2 ;
  format lf2 lf3 sph mu1 mu2 sg1 sg2 6.3 ;
  var n k lf2 lf3 sph mu1 mu2 sg1 sg2 ;

proc print data=main6 ;
  var index g y s e ;

proc means mean ;
  title2 'Means Absolute Deviation of Prediction Error' ;
  var e ;
```

Table 1

Normality Test on Bids and Asks

			Percent Rejected at 10% Level		
Session	Nobs	Action	Informed	Uninformed	Combined
9	30	Bid	0.50	0.53	0.63
		Ask	0.40	0.53	0.67
10	25	Bid	0.68	0.60	0.76
		Ask	0.64	0.40	0.76
11	30	Bid	0.43	0.47	0.63
		Ask	0.30	0.30	0.57
12	30	Bid	0.63	0.53	0.77
		Ask	0.70	0.40	0.77
14	30	Bid	0.43	0.50	0.73
		Ask	0.37	0.37	0.67
15	40	Bid	0.78	0.85	0.95
		Ask	0.93	0.73	0.98
16	40	Bid	0.73	0.68	0.98
		Ask	0.60	0.78	0.88
17	35	Bid	0.63	0.51	0.57
		Ask	0.54	0.57	0.69
All	260	Bid	0.61	0.60	0.77
		Ask	0.57	0.53	0.76

Notes: A Shapiro-Wilk Normality test is performed for each period on the bids and asks. The percent of the times the test rejects Normality at the 10% level is reported for each session.

Table 2

Statistics on Bids and Asks

I. Paired T-test for Same Means

Action	Number of Periods	Statistic	Probability
Bid	260	7.10	<0.01
LogBid	260	5.04	<0.01
Bought	260	-0.11	0.91
Ask	260	-9.19	<0.01
LogAsk	260	-8.03	<0.01
Sold	260	0.00	0.99

II. F-test for Same Variances

Action	Number of Periods	Percent Rejected at 10% Level
Bid	260	0.37
Ask	260	0.37

Notes: In Part I, a paired Student's T-test for same means is performed by taking the informed less the uninformed action for each period within each session. In Part II, an F-test for same variances is performed for each period within each session. The overall percentage of the times the test rejects the null of same variances at the 10% level is reported.

Table 3a

Results of the Mixing Model for Session 16

Informed Group

Variable	N	Mean	Std Dev	Minimum	Maximum
GROUP	20	1.0000000	0	1.0000000	1.0000000
PRICE	20	1.7555000	0.4763510	0.7900000	2.7700000

Uninformed Group

Variable	N	Mean	Std Dev	Minimum	Maximum
GROUP	20	2.0000000	0	2.0000000	2.0000000
PRICE	20	2.0010000	0.2373960	1.4000000	2.1800000

Maximum Likelihood Results

K	LF2	LF3	SPH	MU1	MU2	SG1	SG2
15	12.699	-14.869	0.520	1.668	2.106	0.437	0.070

Notes: Descriptive statistics are provided for the informed and uninformed groups. The maximum likelihood results after 15 iterations show a likelihood value of 12.70 or -14.87 without the constant term. The percentage of observations from the informed group is 52%. The mean and variances from each distribution are also shown.

Table 3b

Details for Period 1 of the Mixing Model for Session 16

OBS	EXP	ACTION	PRICE	GROUP	NGROUP	S	ERROR
1	16	BID	2.00	2	0	0.29129	0.29129
2	16	BID	1.00	1	1	1.00000	0.00000
3	16	BID	1.50	1	1	1.00000	0.00000
4	16	BID	1.60	1	1	1.00000	0.00000
5	16	BID	1.40	2	0	1.00000	1.00000
6	16	BID	1.65	2	0	1.00000	1.00000
7	16	BID	1.45	2	0	1.00000	1.00000
8	16	BID	1.70	1	1	1.00000	0.00000
9	16	BID	1.80	2	0	0.99954	0.99954
10	16	BID	2.22	1	1	0.22460	0.77540
11	16	BID	2.25	1	1	0.36614	0.63386
12	16	BID	1.00	1	1	1.00000	0.00000
13	16	BID	2.00	2	0	0.29129	0.29129
14	16	BID	2.01	2	0	0.24755	0.24755
15	16	BID	0.79	1	1	1.00000	0.00000
16	16	BID	2.00	1	1	0.29129	0.70871
17	16	BID	2.05	2	0	0.14123	0.14123
18	16	BID	2.77	1	1	1.00000	0.00000
19	16	BID	1.76	1	1	0.99997	0.00003
20	16	BID	2.06	2	0	0.12684	0.12684
21	16	BID	2.10	2	0	0.09718	0.09718
22	16	BID	2.11	2	0	0.09494	0.09494
23	16	BID	2.11	1	1	0.09494	0.90506
24	16	BID	1.70	1	1	1.00000	0.00000
25	16	BID	1.29	1	1	1.00000	0.00000
26	16	BID	2.12	2	0	0.09442	0.09442
27	16	BID	2.15	2	0	0.10331	0.10331
28	16	BID	2.13	2	0	0.09559	0.09559
29	16	BID	1.70	1	1	1.00000	0.00000
30	16	BID	2.16	2	0	0.11024	0.11024
31	16	BID	2.16	2	0	0.11024	0.11024
32	16	BID	2.03	1	1	0.18276	0.81724
33	16	BID	2.16	2	0	0.11024	0.11024
34	16	BID	1.87	1	1	0.97801	0.02199
35	16	BID	2.16	2	0	0.11024	0.11024
36	16	BID	2.17	2	0	0.11963	0.11963
37	16	BID	2.18	2	0	0.13194	0.13194
38	16	BID	2.07	1	1	0.11571	0.88429
39	16	BID	2.00	1	1	0.29129	0.70871
40	16	BID	1.75	1	1	0.99998	0.00002

Notes: The predictions of the model are compared with the actual data. The probability of the action belonging to the informed group is given by the variable S. The error is the difference between the predicted probability and the actual state. The mean error is 0.293.

Table 4

Mean Absolute Errors for All Sessions

Session	Nobs	MAD	Nobs	MAD (filtered)
Bid Only				
9	416	.439	218	.404
10	349	.488	236	.498
11	708	.455	543	.437
12	521	.477	376	.459
14	591	.498	340	.529
15	898	.494	645	.462
16	940	.510	469	.500
17	802	.490	483	.498
Ask Only				
9	546	.455	352	.435
10	450	.493	324	.462
11	579	.442	378	.397
12	741	.456	322	.388
14	581	.502	255	.480
15	713	.452	415	.427
16	983	.474	636	.463
17	806	.501	416	.499

Notes: The prediction errors of mixture model are reported for each session based on mean absolute deviations. Since there are only two possible groups, the naive prediction error is .50. The mean absolute deviations are also reported for a model which filtered out predictions where the difference in the predicted mean value were less than 20¢.

Table 5

Comparison of Actual Signals and Estimated Information

Information through Participation Benchmark

		Estimated Means from the Mixing Model					
Session	Price	Optimal	Weighted	Simple	Pooled	MinVar	Max
Bids Only							
9	.343	.350	.366	.382	.344	.467	.370
10	.319	.291	.312	.338	.319	.499	.292
11	.262	.312	.293	.342	.262	.293	.220
12	.298	.277	.275	.363	.299	.443	.265
14	.318	.306	.293	.349	.318	.440	.288
15	.344	.277	.257	.510	.344	.286	.229
16	.266	.225	.241	.278	.266	.266	.202
17	.295	.245	.274	.307	.295	.364	.295
Asks Only							
9	.225	.204	.238	.296	.225	.265	.125
10	.241	.243	.228	.305	.241	.273	.181
11	.137	.126	.124	.284	.137	.380	.175
12	.218	.230	.222	.250	.218	.249	.178
14	.215	.184	.212	.270	.217	.233	.152
15	.319	.340	.311	.586	.320	.285	.133
16	.269	.256	.241	.302	.270	.212	.131
17	.276	.254	.251	.310	.277	.236	.184

Notes: Price along with various estimators are compared to a benchmark which uses the actual signals provided to traders. Each time a trader participates in the market, the trader's signal contributes to the information in the market. The mean absolute deviation for each comparison is reported. The optimal estimator uses the variance approximation discussed in section 3.3, and is applied on a period-by-period basis.

Table 6

Comparison of Actual Signals and Estimated Information

Fully Aggregated Information Benchmark

		Estimated Means from the Mixing Model					
Session	Price	Optimal	Weighted	Simple	Pooled	MinVar	Max
Bids Only							
9	.385	.387	.413	.406	.387	.492	.377
10	.355	.347	.338	.361	.355	.474	.295
11	.272	.312	.288	.339	.272	.336	.205
12	.330	.312	.323	.401	.331	.465	.234
14	.361	.381	.363	.382	.361	.446	.285
15	.360	.274	.291	.488	.342	.298	.224
16	.253	.218	.233	.273	.253	.270	.204
17	.330	.264	.296	.340	.330	.371	.286
Asks Only							
9	.139	.112	.140	.185	.139	.252	.227
10	.172	.156	.213	.287	.172	.428	.254
11	.123	.139	.130	.237	.123	.432	.360
12	.112	.143	.138	.190	.112	.208	.201
14	.155	.121	.164	.238	.156	.326	.217
15	.242	.246	.203	.498	.242	.519	.218
16	.219	.219	.215	.246	.222	.282	.182
17	.200	.181	.186	.248	.201	.243	.183

Notes: Price along with various estimators are compared to a benchmark which uses the actual signals provided to traders. The signal for each trader is aggregated regardless of the trader's participation in the market. The mean absolute deviation for each comparison is reported. The optimal estimator uses the variance approximation discussed in section 3.3, and is applied on a period-by-period basis.