Bayesian Analysis of Information and Net Order Flow in a Learning Model of Pricing

Ron Guido
School of Banking and Finance
University of New South Wales
Sydney NSW 2052 Australia

We develop a microstructure model which seeks to describe the way in which private information is incorporated into price via a bayesian learning process used by market agents. A novel innovation is the description of a dynamic latent process for the information arrival process which can be extracted from observed order flow variables by developing various filtering techniques. We also describe how a boundedly rational learning process by which agents use aggregated trade data to infer this latent process can be used to set prices. The implication of this specification is that it allows us to examine how various microstructure parameters including the probability of informational trading, the spread attributable to information asymmetry and the degree of information asymmetry vary over time. We adopt Bayesian MCMC methods to directly estimate the structural parameters and test the implications of the model. In relation to modelling trading behaviour, net order flow seems to be well described by a process in which the information state variable follows a simple Markovian process. This is achieved in part by modelling the probability of an information event as time varying.

Introduction

Understanding the way in information is incorporated into markets is of great interest to researchers. The first theoretical work in this area include Kyle (1985) and Glosten and Milgrom (1985). Investigating the predictions made by these theoretical models has also been actively pursued by researchers.

Many previous studies on intraday price formation have relied on running reduced form equations to the test the implications of these microstructure models (see Madhavan (1992), Madhavan and Schmidt (1991), Madhavan et al (1997), Glosten and Harris (1988) and Huang and Stoll (1997)). The reason for this is that many underlying processes that drive such microstructure models are latent, and hence unobservable.

This paper instead both develops and tests a model of intraday price formation in a prototypical quote driven market, where order flow acts as a signal of the security's future value. Central to this model is the dynamic structure of the latent information process which uniquely characterise the arrival of trades and the learning process used by market agents to infer the underlying state of nature. This learning process in turn determines the trading and price setting behaviour of these market agents. This study pays special attention to the modelling and estimation of this dynamic process. A methodology is developed that uses observed prices and order flow to efficiently filter the latent variables and directly estimate and test the parameters of the structural model posited. This methodology is based on Markov Chain Monte Carlo strategies, namely via Gibbs sampling, to obtain the posterior densities of the structural parameters which can then be used to estimate and test the model. This allows us to make direct comparisons among various competing models of order flow through the implementation of Bayesian hypothesis testing and model selection techniques (see Han and Carlin, (2001), Kass and Raftery (1995) and Chib (1995)).

The analysis examines the information content of trade direction and how net order flow or trade imbalance coupled with trading activity may be used by market agents to ascertain the underlying information state. A key feature of the model describes the way in which market agents infer the underlying state of the market. Unlike many previous market microstructure models where agents update beliefs following each order arrival, the concept of bounded rational learning is introduced, whereby the market agent, namely the market maker, waits for a period of time observing the history and sequence of trades before updating his beliefs. This idea can be motivated by the fact that it can be too costly to efficiently update beliefs following each trade, particularly when it may be more informative to observe the trading activity in any given period of time. This will clearly have an impact on the relation between order flow and prices as represented by the quotes set by the market maker.

Furthermore, few theoretical microstructure models developed thus far have considered the differential roles of trading frequency and net order flow in explaining the evolution of information and prices; as such we make a significant contribution to the existing body of research in this area of market microstructure and information theory.

The primary motivation in adopting the MCMC methodology is the conceptual and computational simplicity afforded by these techniques when dealing with latent variables. The conceptual simplicities arise by treating the unobserved data in the same manner as the unobserved parameters, while the computational simplicity results from characterising analytically intractable densities using simulation. Excellent introductions to MCMC analysis include Casella and George (1992), Chib and Greenberg (1996) and Tanner (1996).

The paper is organized as follows. The following section (Section 1) considers the theoretical foundations of the order flow model, the dynamic structure of information flow and the expectation formation mechanism used to set prices. Section 2 explains in detail the order flow specification and the MCMC techniques necessary to estimate the structural parameters and the stochastic process governing information flow. In sections 3 and 4, the price formation models and the MCMC procedures necessary to estimate the structural parameters are examined. Section 5 presents the results, and conclusions are drawn in section 6.

1. A Model of Information Flow, Net Order Flow and Leaning Behaviour in a Quote Driven Market

Dynamics Governing Information Flow and Trading Structure

In this section we develop a theory of the stochastic processes governing information flow and trade structure by modelling order flow.

As noted in the seminal works by Glosten and Milgrom (1985) and Easely and O'Hara (1992) there is difficulty in specifying the mechanics of the order arrival process as it would require specifying both individual trader behaviour and any frictions that might be present in the trading mechanism. This difficulty led both sets of research to adopt the convention of an exogenous order arrival process, whose parameters are static in nature. By better defining this process within a dynamic setting, we are able to model the evolution of state probabilities that are used by market agents to form their expectations about future order flow and the underlying value of the security. The formation of these expectations and their computation allows us to accurately construct the price quotations as set by the specialist.

We begin by considering a prototypical trading structure as adopted by several market microstructure studies (See Kyle 1985, Glosten and Milgrom (1985) and Easly and O'Hara (1992)). Suppose that S_t is an unobserved discrete random variable with domain: {1, 2, 3}. This variable represents a private signal observable to certain market traders as to the underlying value of the security. Specifically, S_t represents whether at time t, no information event has occurred ($S_t = 1$), an information event with a high signal (a positive news event) has occurred ($S_t = 2$), or an information event with a low signal (a negative news event) has occurred ($S_t = 3$). This private signal will only be observed by a certain fraction of traders who are considered informed. These traders will act upon this information by deciding whether to buy, sell or not to trade conditional on the signal observed. The remaining fraction of traders like the market maker are considered uninformed and do not observe the signal. These traders trade mainly for liquidity based reasons; this ensures that no-trade equilibria are avoided.

Unlike previous studies, we assume S_t evolves according to a first-order Markov process such that the probability of a certain state of nature occurring is time varying, with its transition probability matrix given by

$$\Pi = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{1}$$

where p_{ij} = Pr(S_t = j | $S_{t\text{-}1}$ = i) and $\sum_{j=1}^{3} p_{ij} = 1$ for all i. Given this Markov scheme, we describe the conditional distribution of a (vector) process, X_t, as

$$X_t \mid \Phi_{t-1} \sim \begin{cases} f(X_t \mid s_t = 1, \Phi_{t-1}) \\ f(X_t \mid s_t = 2, \Phi_{t-1}) \\ f(X_t \mid s_t = 3, \Phi_{t-1}) \end{cases} \tag{2}$$

where Φ_{t-1} represents the information set thus available to economic agents at time t-1. Φ_{t-1} can be thought of as the history of past trading outcomes $X^{t\text{ -}1}$ = ($X_{t\text{-}1}$, $X_{t\text{-}2}$,…, X_1), but can be generalised to include the history of all publicly available information. We postulate that economic agents adopt at least two measures to define the trade outcome variable; net order flow, Q_t, and trade activity, n_t, for the period t such that X_t = (Q_t , n_t). Net order flow, also known as net order imbalance, represents the difference between buyer initiated and seller initiated order flow over a given period, whereas trade activity represents the number of trades occurring over a given period. For the purposes of this study we consider the measure of order flow based upon a trade direction variable which equals 1 when a trade is buyer initiated and –1 when it is seller initiated.

Specification of the Learning Process used by market agents

In a vein similar to the market microstructure models that use the trade process as a learning tool (Easley and O’Hara, 1987, 1992 and Easely, Kiefer and O’Hara, 1997(a) and (b)), we utilise the probabilistic structure of the order arrival process to highlight how the market maker can infer the underlying state. The market maker is a Bayesian who knows the structure of the market, or more precisely the structural parameters governing the information and order flow process. Following a trading outcome the market maker updates his or her posterior sate probabilities using Bayes rule. Given that the probabilistic structure is one defined in terms of the state switching probabilities, we adopt the novel approach used by Gray (1996) and Hamilton (1994) to translate a model expressed in terms of state switching probabilities, Pr(S_t | $S_{t\text{-}1}$), to one in terms of regime probabilities, Pr($S_t \mid \Phi_{t-1}$). As a result we are able to construct the market maker’s posterior probability of each state conditional only on the history of order flow.

To represent the model in terms of these regime probabilities, define $p_{t/t\text{-}1}$ = [p_{1t} p_{2t} p_{3t}] where p_{it} = Pr(S_t = $i \mid \Phi_{t-1}$). Similarly, we define $p_{t,t}$ = [p_{1tt} p_{2tt} p_{3tt}] where $p_{it,t}$ = Pr(S_t = i $\mid \Phi_t$). By the first order Markov property of S_t, we can write

$$p_{t/t-1} = \Pi' p_{t-1/t-1}.$$

Defining the vector of conditional likelihoods as

$$f_t = \begin{bmatrix} f(X_t \mid s_t = 1, \Phi_{t-1}) \\ f(X_t \mid s_t = 2, \Phi_{t-1}) \\ f(X_t \mid s_t = 3, \Phi_{t-1}) \end{bmatrix} \tag{3}$$

it can be shown (see Hamilton, 1990 and1994) that,

$$p_t \equiv \Pr(s_t \mid \Phi_{t\text{-}1}) = \Pi' \left[\frac{f_{t-1} \otimes p_{t-1}}{\iota'(f_{t-1} \otimes p_{t-1})} \right] \tag{4}$$

by application of Bayes Rule. We can then utilise a recursive procedure in which to construct the likelihood function and hence conveniently compute expected value quantities. (See Appendix for details).

2. Modelling Order Flow and its Bayesian Analysis

In this section we construct a model of order flow in order to better describe its dynamics, as well as develop Markov Chain Monte Carlo (MCMC) strategies to permit a Bayesian analysis of the model.

The idea developed in this study is that market participants observe the aggregate of trading behaviour in order to infer the underlying state of nature. By doing so, both trading activity as reflected by the number of trades and the net order imbalance in any given period become informative measures used by market agents to learn the state of the market.

The models to be considered all belong to a general class where the distribution of observations depends upon a latent Markovian switching process in a discrete state space. A MCMC scheme is developed in order to construct a Bayesian analysis of the parameters and evaluate the performance of the various models considered. As with all mixture models, they potentially suffer from unidentifiability, in which case steps must be taken in order to directly estimate the state specific parameters of the model. This will be also considered in detail in later sections.

Net Order Flow, Trade Activity and Bounded Rationality

The weakness of most traditional sequential trade models, is that for the most part the timing convention relates to the ordinal sequence of trades and not periodic time intervals. These models are unable to characterize changes in the degree of trading activity which in itself would be informative to market participants. While this was a problem also considered by Easely and O'Hara (1992), they addressed this issue by modelling the number of no–trade outcomes and showing how the absence of trade could be used to infer the prevailing information state. They also described how the cumulative number of trades could act as a signal as to the existence of information, while the cumulative order imbalance (as reflected by the market maker's inventory) could provide a signal of the direction of any new information. To adequately address these issues we move beyond a transaction level analysis and aggregate trades thereby allowing an analysis of the variation in trading activity.

By modelling order flow in this fashion, we postulate a mechanism by which the market-maker uses trade sequence as well as the amount of trading in a given time interval to determine the state of nature. While quotes set by the market maker are still formed on a transaction-by-transaction basis, the state probabilities used to form their expectations about the underlying value of the security are updated on a periodic basis rather than following each transaction. This approach assumes a form of *bounded rationality* implicit in the market maker's behaviour; in that it would be costly to update beliefs upon the arrival of every piece of new information (individual trades), and would prefer to observe a sequence of trades to infer the degree of trading activity. This section develops simple models of trading activity and shows how trading activity can be used as an informative variable by the market maker.

To do this requires certain assumptions. In determining the appropriate time interval in which market agents use to aggregate trades, we assume that the state as described by the latent variable, S_t, prevails for that entire period and then may change according to its transition probability matrix defined by Π. We assume that the transition of S_t evolves independently over time. However, informed traders condition on the state signal they observe at the commencement of the period thus implying that trade activity and hence net order flow is state dependent. We would expect that an information event/state would be associated with a high degree of trading activity over that period; with the sign of the trades determining direction of the information state (high or low).

Net Order Flow and Trading Activity in a Trade Direction Model

In order to model aggregate volume, let [t-1, t] represent a fixed time interval. Denote the realised trade variable, $x_{\tau,t}$, to represent the τth trade occurring during the time interval [t-1, t]. The random variable, n_t, represents the number trade outcomes in the interval [t-1, t].

Based upon these variables, we construct, Q_t, the net order flow for the current period t. By definition, not order follow can be modelled as

$$Q_t = \sum_{\tau=1}^{n_t} x_{\tau,t} \; . \tag{5}$$

This variable, together with the trade frequency variable, n_t, is used by the market maker to learn the prevailing state of nature. In order to describe its dynamic properties, we model Q_t as a state contingent compound process. To do this we first need to model x_t and n_t.

Given that n_t represents the number of trades occurring in a given time interval, we can model the trade frequency variable as a state contingent poisson process: $n_t \mid s_t = i, \Phi_{t-1} \sim P(\vartheta_i)$. Hence n_t can be described as a mixture of poisson distributions indexed by parameter ϑ_i:

$$n_t \mid \Phi_{t-1} \sim \begin{cases} P(\vartheta_1) & \text{if } \; s_t = 1 \\ P(\vartheta_2) & \text{if } \; s_t = 2 \\ P(\vartheta_3) & \text{if } \; s_t = 3, \end{cases}$$

where

$$\Pr(\, n_t \mid \Phi_{t-1}, s_t = i) = \frac{\vartheta_i^{n_t} \exp(-\vartheta_i)}{n_t!} \, . \tag{6}$$

This specification accommodates several features of the microstructure rather well. Given State 1 represents the no information state, trading during this period should be non-informational or liquidity based. Thus the parameter, ϑ_1 captures the trading activity generated by liquidity traders; the difference between ϑ_2 and ϑ_1, reflects the trading activity generated by informed traders during positive states, and the difference between ϑ_3 and ϑ_1 reflects the trading activity generated by informed traders during negative states.

The trade indicator variable, $x_{\tau,t}$, is modelled as a discretely distributed binary random variable dependent on S_t and available information, Φ_{t-1}: $x_{\tau,t} \mid S_t = i, \Phi_{t-1}$ which is restricted to take on values of 1 (buyer initiated trade) and –1 (seller initiated trade). The conditional distribution of $x_{\tau,t}$ can be written as

$$x_{\tau,t} \mid \Phi_{t-1} \sim \begin{cases} p(x_{\tau,t} \mid s_t = 1, \Phi_{t-1}) & \text{if } \; s_t = 1 \\ p(x_{\tau,t} \mid s_t = 2, \Phi_{t-1}) & \text{if } \; s_t = 2 \\ p(x_{\tau,t} \mid s_t = 3, \Phi_{t-1}) & \text{if } \; s_t = 3, \end{cases}$$

where the state contingent likelihoods are defined as

$$p(x_{\tau,t} \mid s_t = i, \Phi_{t\text{-}1}) \; = \; q_{i1} \boldsymbol{I}\{x_{\tau,t} = 1\} + q_{i2} \boldsymbol{I}\{x_{\tau,t} = -1\}, \tag{7}$$

where $\sum_{j=1}^{2} q_{ij} = 1$. The structural parameters q_{ij}, can be interpreted as the parameters which translate trader's latent demands into observed order flow. Specifically q_{i1} is be interpreted as the probability that the next trade is buyer initiated given that the prevailing state, S_t, is equal to i. A-priori, these parameters will vary across states to reflect the behaviour of informed and uninformed traders. For example, we would expect the probability of a buy, q_{i1}, to be larger in the positive event state (S_t =2) than in other states particularly during an negative information state to reflect the increased presence of informed trading activity.

Having described the conditional distributions for n_t and $x_{\tau,t}$, we can construct the conditional distribution for Q_t. We wish to determine and ultimately estimate the parameters governing the process, Q_t (denoted by θ). The aim then is to obtain the density function (likelihood) of the observed random variable Q_t, where this value represents the net order flow for period t. Given that n_t and $x_{\tau,t}$ are both random variables and that $Q_t = \sum_{\tau=1}^{n_t} x_{\tau,t}$, the likelihood for Q_t (suppressing Φ_{t-1} for notational convenience) is given by the following compound process:

$p(Q_t \mid S_t, \theta) = \sum_{j=0}^{\infty} \left\{ f(Q_t \mid j,\, s_t) \frac{e^{-\vartheta} \vartheta^j}{j!} \right\}$, where $f(Q_t \mid j, S_t)$ represents the conditional probability distribution function of Q_t given $n_t = j$.

This likelihood is difficult to use in order to sample the full conditional densities of the parameters, since the density is an infinite sum. One approach involves the application of Bessel functions and considering the formulation as a simple randomization problem when x is discrete (See Feller, Volume II, 1966). However, given that we require the joint density for the trade outcome process $X_t = (Q_t \;\; n_t)$ in order to construct the learning rule for the market maker, we consider an alternative approach that simplifies matters considerably. Given that we observe Q_t and n_t contemporaneously, we can model their joint distribution. Noting that $p(Q_t, n_t \mid s_t, \theta) = p(Q_t \mid n_t, s_t, \theta) \,.\, p(n_t \mid s_t, \theta)$, it is easy to establish the joint likelihood function given knowledge of the probability density function, $p(Q_t \mid n_t, s_t, \theta)$. It can be shown that this density is given by

$$p(Q_t \mid n_t\,,\, s_t = i) \;=\; \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{\,Q_t + v_t} q_i^{\,v_t} \tag{8}$$

where $n_t - Q_t = 2v_t$. (See Appendix for details). Hence the joint likelihood function can be written as

$$f_t \equiv p(Q_t\,,\, n_t \mid s_t = i) \;=\; \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{\,Q_t + v_t} q_i^{\,v_t} \,.\, \frac{e^{-\vartheta_i} \vartheta_i^{n_t}}{n_t!} \,. \tag{9}$$

This expression forms the basis of the likelihood function defined in (3) which is used by the market maker to form the posterior probabilities of each state as determined by the updating equation in (4).

Bayesian Analysis of Order Flow Model

In this section we develop the MCMC techniques necessary to perform Bayesian inference for the order flow models developed thus far. Bayesian Analysis by contrast to classical methods, treats the parameter vector governing the model as a random variable and so a primary aim in Bayesian inference is to derive, or simulate iterates from, the joint posterior distribution of the parameter vector, θ. MCMC methods, namely via the Gibbs sampler, allows us to sample from what can often be analytically intractable densities by constructing a Markov chain on a general state space such that its limiting distribution is the joint distribution of interest. Often, the ability to perform inference using classical procedures such as maximum likelihood are constrained due to computational difficulties and thus leads to the analysis of simpler reduced form models rather than the structural models whose parameters are directly interpretable. As such the primary motivation in adopting the MCMC methodology is the conceptual and computational simplicity afforded by these techniques particularly when dealing with models that contain latent variables (see Gelfand and Smith, 1990) and the ability directly describe the structural parameters of a model.

The order flow models considered here all belong to a general class where the distribution of observations depends upon a latent Markovian switching process on a discrete state space (see Albert and Chib (1993), Carter and Kohn (1994) and Chib (1995)). For the purposes of the current model specifications, we define the following quantities: let θ represent the parameter vector for the model in question, $X^t = (X_1,...,X_t)$ be the observable vector of data up to time *t* where we define $X_t = (Q_t,\ n_t)'$, and $S^t = (S_1,...,S_t)$ represent the vector of the latent state variables up to time *t*. In this context we seek to derive the joint posterior of the parameter vector and the latent state vector given the set of observable data: $f(S,\ \theta|X)$.

Although analytically intractable, by constructing a Gibbs sampler for the problem we can obtain draws of θ and *S* which can be viewed as being drawn form the joint density of interest. By firstly augmenting the parameter vector by the latent vector *S*, the decomposition of the joint posterior density according to Bayes theorem, $f(S,\ \theta|X) \propto f(X\,|\,S,\ \theta)f(S|\ \theta)f(\theta)$, leads to the following algorithm:

1. Initialise θ
2. Sample S from $f(S|\theta)$.
3. Sample θ from $f(\theta\,|S, X)$.
4. Repeat steps 2 and 3.

Under mild regularity conditions, the iterates generated from this sampling algorithm will converge to their invariant target distribution. Given a sufficiently large number of draws, the parameters' marginal posterior distributions can be constructed. Furthermore, by averaging subsets of these simulations Bayesian estimators of the parameters can also be formed. For details, refer to Casella and George (1992), Tanner (1996), and Chib and Greenberg (1996).

As noted by several authors (see Stephens, 2000), Bayesian analysis of mixture models potentially suffer from parameter unidentifiability, due to the invariance and symmetry of the likelihood function, in which case steps must be taken in order to directly estimate the state specific parameters governing the model. This will be also considered in detail in this section.

Obtaining the Bayesian estimators for the parameters governing the net order flow models entails sampling from the set of full conditional posterior distributions. Sampling from these full conditional distributions form the basis for the Gibbs Sampler which leads to the generation of iterates from the joint distribution of the parameters governing the models and are derived in the following sections.

We first treat a discussion of sampling the latent state vector *S*, and the parameters governing the transition probability matrix, Π. Having done so we consider the parameters governing each order flow model in turn by deriving their full conditional distributions.

Sampling S^n (S_t, t = 1,…, n)

We can prescribe the following block-sampling scheme to generate the state variable S^n. Suppressing θ for notational convenience our aim is to generate S^n from the distribution Pr($S^n\ |\ X_n$). We first note the decomposition

$$\Pr(S^n \mid X_n) = \Pr(S_n \mid X_n) \prod_{t=1}^{n-1} \Pr(S_t \mid X_t, S_{t+1}) \tag{10}$$

Following Carter and Kohn (1994) and Chib (1995), we construct an algorithm which samples the vector S^n as a block using the joint conditional distribution, Pr($S^n | X_n$), rather than from the set of individual full conditional distributions $\Pr(S_t|Y_n,S_{-t})$. Since the process $\{S\}_t$ is Markov, and therefore correlated, such blocking will lead to faster convergence to the posterior distribution, and is therefore preferred to the single move sampling. The algorithm first generates S_n from Pr($S_n | X_n$) and then for t=n-1, ,1 successively generates S_t from Pr($S_t | X_t$, S_{t+1}) which from Bayes theorem, for t=n-1,…,1 can be computed from :

$$\Pr(S_t | X_n , \mathrm{S}_{t+1}) = \frac{\Pr(S_{t+1} \mid S_t)\Pr(S_t \mid X_t)}{\Pr(S_{t+1} \mid X_t)}. \tag{11}$$

The discrete filter to evaluate Pr(S_t | X_t) can then be described as a set of recursive equations used generate samples from Pr(S^n | X_n). Details of this algorithm can be found in the appendix.

Sampling the Transition Probability Matrix

Having generated the latent state variable, S, it is a relatively straightforward task to sample from the full conditional distributions of the parameters that form the transition probability matrix, Π.

Let $\Pi_i \equiv (p_{i1},..,p_{iK})$ represent the ith row of Π; the vector of state transition probabilities given $S_t = i$. By construction, these probabilities must sum to unity. The full conditional distribution for Π_i can be expressed as:

Pr(Π_i |X_n, $S^n, \theta_{/\Pi_i}$) = Pr(Π_i | S^n), since Π_i is independent of X_n and $\theta_{-\Pi_i}$ given S^n. By Bayes Rule then Pr(Π_i | S^n) $\propto$ Pr(S^n |Π_i)Pr(Π_i).

Given that S_t evolves according to a first order Markov process, the joint likelihood for S^n can be expressed as a Dirichlet process. By adopting conjugate priors for Pr(Π_i), the posterior denisty too will be a dirichlet, and so that parameters for Π_i can be jointly sampled from the following Dirichlet distribution :

$$\Pi_i \mid S^n \sim \mathrm{Dir}(d_{i1}, d_{i2},\ldots, d_{ik}) \tag{12}$$

where $d_{ij} = n_{ij} + u_{ij}$, n_{ij} represents the number of transitions from state i to state j : $n_{ij} = \sum_{t=2}^{n} \boldsymbol{I}(s_{t-1} = i)\boldsymbol{I}(s_t = j)$ and u_{ij} are the hyperparameters of the Dirichlet prior. See Appendix for details.

Sampling the Parameters Governing Trade Direction Based Model of Net Order Flow

Given the trade direction model developed earlier, the parameter vector can be defined as $\theta = (q, \vartheta, vec(\Pi))$. Given the joint sampling density or conditional likelihood for X_t, we can construct joint density of the vector of observations conditional on state i. By defining I_{it} to represent an indicator function that equals one when $S_t = i$, and zero otherwise, we can express the conditional likelihood as:

$$p(X_T \mid S^T, q_i, \vartheta_i, \Pi) = p(X_T \mid S^T, q_i, \vartheta_i)$$

$$= \prod_{t=1}^{T} \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{Q_t+v_t} q_i^{v_t} I_{it} \frac{e^{-\vartheta_i} \vartheta_i^{n_t}}{n_t!}$$

$$= \left(\prod_{t=1}^{T} n_t! \right)^{-1} \prod_{t=1}^{T} \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{\sum_{t=1}^{T}(Q_t + v_t)} q_i^{\sum_{t=1}^{T} v_t} \; e^{-T\vartheta} \vartheta_i^{\sum_{t=1}^{T} n_t} \; I_{it} \,,$$

$$\propto p_i^{\sum_{t=1}^{T}(Q_t + v_t)} q_i^{\sum_{t=1}^{T} v_t} \; e^{-T\vartheta} \vartheta_i^{\sum_{t=1}^{T} n_t} \; I_{it} \,. \tag{13}$$

As shown in the appendix, using conjugate priors we are able to construct the full conditional densities of the trade direction parameters to which we sample from in the MCMC algorithm:

$$\Pi_i \mid S^n \sim \mathrm{Dir}(\delta_{i1}, \delta_{i2},\ldots, \delta_{ik}) \tag{14}$$

$$q_i \mid X^T, S_T, \theta_{/q_i} \sim \text{ Beta } (b_i + \textstyle\sum_{t=1}^{T}(Q_t + v_t) \,, s_i + \sum_{t=1}^{T} v_t) \tag{15}$$

$$\vartheta_i \mid X^T, S_T, \theta_{/\vartheta_i} \sim \text{Gamma } (\textstyle\sum_{t=1}^{T} n_t + c_i, T + d_i) \tag{16}$$

Modelling Issues in Bayesian Analysis: Label Switching in Mixture Models

When conducting Bayesian analysis of mixture models, parameter estimation can be complicated by the inability of the Markov chain to generate parameter iterates that belong solely to a single mixture component. This so-called label switching problem generally arises when taking a Bayesian approach to parameter estimation within mixture models. The problem has been identified by several authors, including Diebolt and Robert (1994), Richardson and Green (1997), and Fruhwirth-Schnatter (2001) and Celeux et al (2000). The problem arises due to the fact the likelihood and hence posterior distribution of the model parameters under diffuse priors are symmetric and hence invariant under relabelling of the mixture components. The MCMC sampling thus produces posterior distributions that are multi-modal and highly symmetric, rendering useless inference methods that summarise the parameters by their marginal distributions (e.g. by computing the posterior mean and mode).

Several attempts have been made to remove label switching, the most popular being those which impose artificial identifiability constraints (see Richardson and Green, 1997). Yet this approach does not always provide a satisfactory solution as noted by Celeux et al (2000), and Fruwirth-Schnatter (2001), particularly when there may be no prior knowledge as to how to label the parameters.

The most promising approach however has been that developed by Stephens (2000) which attempts to relabel the iterates for each parameter by selecting the relabelling that minimises the posterior expected loss for a certain class of loss functions. An online algorithm has been adopted in this study, which attempts to relabel the parameters following each sweep of the Gibbs Sampler. All results reported in this study have been successfully relabelled using this algorithm. For details, the reader is referred to Stephens (2000).

3. Bid/Ask Price Dynamics in a Model with Bounded Rational Learning

In this section, we construct a model of asset price formation that synthesises the concepts of bounded rational learning behaviour and order flow dynamics. Central to this model is an expectation formation mechanism used by the market maker to determine the underlying value of the security. The adverse selection parameter or informational asymmetry with the system is modelled as a time varying, state dependent process, which allows the adverse selection component of the bid ask spread set by the market maker to reflect his or her time varying state uncertainty. The market maker’s event uncertainty is summarised by the state probabilities that are formed using the bounded rational learning procedure described in earlier sections. Bounded rationality is a result of the market maker updating these state probabilities using aggregated trades or net order flow to infer the degree of trading activity, which acts as a signal to the market of whether an information event has occurred. In so doing we construct a positive model that describes the relationship between order flow, learning and prices.

MCMC methods are again used to carry out Bayesian inference as they are particularly useful in dealing with the latent variable structure inherent to the model. While previous studies (see Manrique and Shephard, 1997, and Hasbrouck, 1999) have utilised Bayesian methods to investigate various microstructure effects, our study goes further by investigating the stochastic process of the latent variables acting upon prices and order flow. In so doing, issues such as those pertaining to persistence in the volatility of prices may be explained by the process in which the market learns about the prevailing latent information regime. This learning process coupled with the dynamics of the state variables can then be used to explain the observed behaviour of prices and volume.

Model Development

In order to develop the model, we first consider a simple a sequential trade model of price formation that employs several elements of the framework adopted by Madhavan, Richardson and Roomans, 1997 (MRR, 1997), and is considered in detail below.

Consider a quote driven market for a security with a competitive market maker who provides liquidity and permits continuous trading by overcoming the asychnronous timing of investor orders. These liquidity providers quote two prices: the bid price, at which they will buy securities and the ask price at which they will sell securities. The market maker posts bid and ask prices that are ex-post rational, so that the pre-trade ask (bid) price at ordinal time *t* represents the expected value of the security conditional on the history of all available information, Φ_{t-1}, and a buyer (seller) initiated trade. The observed spread in the quotations also reflects the market maker's compensation for liquidity provision. The formation of these conditional expectations and its revision are dependent on the specialist's belief of the information content of incoming trades and any public information that arrives to the market. Given we are modelling time as an ordinal sequence, we re-define the order flow variable occurring at instant *t* to be x_t.

To model the revision in beliefs, denote V_t as the true underlying value of the security and let m_t denote the post-trade expected value of a risky security conditional on public information and order flow at time *t*:

$$m_t \equiv E(V_t) = m_{t-1} + \lambda(x_t - E[x_t \mid \Phi_{t-1}]) + \varepsilon_t \qquad (17)$$

This equation describes how the revision in the expected value of the security arises from two sources; new public information and order flow. New public information announcements can cause revision in beliefs without trading volume. We let ε_t represent the innovation in beliefs occurring between *t*-1 and *t* due to the arrival of public information and we assume that $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. The reaction by the market maker to private information is manifested by a revision in beliefs resulting from order flow. As in MRR (1997) as well as Hasbrouck (1991), we assume that the revision in beliefs due to order flow is given by $\lambda(x_t - E[x_t \mid \Phi_{t-1}])$, where $(x_t - E[x_t \mid \Phi_{t-1}])$ represents the market maker's unexpected amount of order flow (or order flow innovation) given his information set Φ_{t-1}, and λ represents the price impact of the innovation or the degree of information asymmetry present in the market.

Dependent upon the specification as described in earlier sections, the order flow variable, x_t, can be defined in several ways. In a *trade direction* model the market considers the flow of buys and sells in order to learn the state of the market. At this stage however, the current model does not accommodate the market maker's ability to observe the aggregate trade outcome (net order flow), whereby trading frequency or the degree of trading activity in any one period can be used to infer the underlying state of nature. The information set available to the market maker and to the public, Φ_{t-1}, can now be thought of as the history of past trading outcomes (x_{t-1}, x_{t-2},…, x_1), and prices, but can be generalized to include the history of all publicly available information.

We denote the bid and ask prices set by the market maker as p_t^b and p_t^a respectively. In establishing the ask and bid price, the market maker is subject to the non-negative costs, ϕ^a, ϕ^b per unit of trade, which will be reflected in the quotes set. In the absence of other costs or frictions, the market maker will quote a bid price of m_t - ϕ^b and an ask price of $m_t + \phi^a$. Given ex-post rationality, quotes at time t are set prior to observing the incoming trade, x_t. In a model where trades are characterized as a buy or sell (x_t = 1 or -1), we can write the ask and bid prices as:

$$
\begin{aligned}
p_t^a &= E[V_t \mid x_t = 1] + \phi^a + \xi_t^a \\
&= m_{t-1} + \lambda(1 - E[x_t \mid \Phi_{t-1}]) + \phi^a + \varepsilon_t + \xi_t^a, \qquad (18)\\
p_t^b &= E[V_t \mid x_t = -1] - \phi^b + \xi_t^b \\
&= m_{t-1} - \lambda(1 + E[x_t \mid \Phi_{t-1}]) - \phi^b + \varepsilon_t + \xi_t^b. \qquad (19)
\end{aligned}
$$

The terms ξ_t^a and ξ_t^b are independent and identically distributed random variables distributed such that $\xi_t^a \sim N(0, \sigma^2_{\xi^a})$, and $\xi_t^b \sim N(0, \sigma^2_{\xi^b})$. These terms are included to account for the effect of stochastic rounding errors induced by price discreteness in the bid and ask price.

In the model proposed by MRR (1997), the latent variable, m_t is substituted out by assuming a simple specification for the temporal behaviour for order flow. This results in a reduced form equation that can then be estimated. This has been a common practice in several microstructure studies including Huang and Stoll (1997), Glosten and Harris (1988) and Madhavan and Schmidt (1991). As the focus of this study is on the mechanism in which information is incorporated in the trading process, we must develop an approach that allows us to explicitly model the latent information structure that in turn determines the stochastic process, x_t. Observing the dynamics of x_t, and its interaction with the posted bid and ask prices allow us as econometricians to infer the properties of the latent or state variables that drive the model.

Given that the trade variable x_t is modelled as a function of the latent information state variable, S_t, we can investigate how information is incorporated into prices by modelling learning behaviour of the market maker. In the sequential trade models of Glosten and Milgrom (1985) and Easley and O'Hara (1987, 1992), trade can arise from uniformed and or informed traders. Trade takes place in a sequential fashion with traders arriving to the market according to a probabilistic process. With the arrival of trades, the market maker uses Bayesian learning to form his expectation of next period's order flow $E[x_t \mid \Phi_{t-1}]$. This expectation is then fed into the ex-post rational quote setting process.

A General Model of Price Expectation Formation under Event Uncertainty

The expectation formation mechanism characterised in (17) assumes a form such that the bid ask spread attributable to private information effects is constant. A new formulation is developed whereby the bid-ask spread and thus the degree of information asymmetry is both time varying and state dependent.

In this model we generalise the expectation formation mechanism of the market maker. If the market maker knew the state S_t with certainty then it is plausible to assume that the trade impact parameter, λ, would be different across these states. A-priori, if S_t =1, then λ should be set to zero or a value commensurate with volume having a small impact on prices. When there is no information event, order flow should convey no new information and hence should have a minimal impact on price. Conversely if an information event occurs, the impact of a buy should be different to that of a sell depending on the direction of the signal; buys should only have an impact on prices if a positive information event is known with certainty, and sells should have an impact if a negative information event is known with certainty.

This requires modelling the evolution of λ and more generally the error correction mechanism, $\lambda(x_t - E[x_t \mid \Phi_{t-1}])$, over time and across states. Given event uncertainty and the learning behaviour of the market maker, we postulate the market maker forms an expectation of the appropriate price impact of the order flow innovation given his posterior probabilities as to the prevailing state given the trade outcome, $p_{it/t} = \Pr(S_t = i \mid \Phi_t)$. The post trade expected value of the stock can then be constructed.

The idea rests upon treating λ as a conditionally monotonic function of order flow. However given state uncertainty, the unconditional behaviour of λ will display a non-monotonic relation with order flow. Let λ_{it} represent the price impact parameter that is determined by the market maker at time t if state i were known with certainty. Under each state, λ_{it} has a specific functional form which is dependent on the order flow variable:

$$\lambda_{it} = \begin{cases} g_1(x_t) & \text{if} \quad S_t = 1 \\ g_2(x_t) & \text{if} \quad S_t = 2 \\ g_3(x_t) & \text{if} \quad S_t = 3 \end{cases} \tag{20}$$

If order flow was modelled as a trade indicator variable, we can prescribe the following functional form:

$$\lambda_{it} = \begin{cases} \lambda_1 & \text{if} \quad S_t = 1 \\ \lambda_1 + \lambda_2 I(x_t = 1) & \text{if} \quad S_t = 2 \\ \lambda_1 + \lambda_3 I(x_t = -1) & \text{if} \quad S_t = 3 \end{cases} \tag{21}$$

Many alternative specifications are however possible. λ_1 represents the non-informational impact of order flow. Since this value prevails under no-information states, we would expect this value to equal zero: a value greater than zero would reflect the effect of liquidity on price setting. Under state S_t = 2, only buy trades are strictly informative; as such λ_2 reflects the additional informational impact of buy trades occurring; sells occurring when S_t = 2 are non-informative or liquidity based and so their price impact is reflected by the value λ_1. Conversely under state 3, sells are strictly informative given the negative news state, and the price impact is reflected in λ_3; λ_1 reflects the price impact of non-informational buyer initiated trades occurring in a negative news state. We can specify the following expectation formation mechanism as follows:

$$m_t = m_{t-1} + \sum_{i=1}^{3} \left\{ \lambda_{it} (x_t - E[x_t \mid \Phi_{t-1}, s_t = i]) p_{it/t} \right\} + \varepsilon_t \tag{22}$$

where $m_t \equiv E[V_t \mid \Phi_t] = E[V_t \mid \Phi_{t-1}, x_t]$. The above process can be a considered a generalisation of earlier mechanisms considered by MRR (1997) in that (33) reduces to their specification when one assumes $\lambda_{it} = \lambda$.

The term $\lambda_{CE} \equiv \sum_i \left\{ \lambda_{it} (x_t - E[x_t \mid \Phi_{t-1}, s_t = i]) p_{it/t} \right\}$ in equation (22) also has a straightforward interpretation; λ_{CE} represents the weighted average of errors that would prevail under each state with the weights being the state probabilities. Under the assumption of risk neutrality, this value reflects the certainty equivalent of the adverse selection component of prices: $\lambda(x_t - E[x_t \mid \Phi_{t-1}])$.

Quote Setting Under Various Trade Model Structures

The quote-setting functions expressed thus far have only considered trade direction as the order flow variable. We now generalize the model of MRR (1997) to allow quote setting behaviour to be contingent either upon trade direction or order size. We also incorporate the generalized expectation formation mechanism described above, as well the bounded rational learning rule used by the maker, into the quote setting functions.

To generalise the quote setting functions defined in MRR (1997) we allow the expectations mechanism used to set prices to be a function of signed trade size as opposed trade direction alone. This corresponds with several theoretical microstructure models including Kyle (1985) and Madhavan and Schmidt (1991) whereby the prices set by the market maker can be contingent on order size. In these cases, the market maker provides the trader with a price schedule in which the trader can then select whether and how much to trade.

In order to avoid specifying full pricing functions for the market maker, Lee, Mucklow and Ready (1993) and Kavajecz (1999) note that the bid and ask quotes represent only one dimension of the quotation provided by the market maker. In markets such as the NYSE, complete quotes consist of the best price for the bid and ask, as well as the number of shares available at these prices (the depths). Hence we can interpret the actual quotes observed in a market by treating the ordered pairs (ask price, p_t^a, and depth at the ask, x_t^a) and (bid price, p_t^b, and depth at the bid, x_t^b) as two points on the ask and bid pricing functions:

$$p_t^a(x_t^a) = E[V_t \mid x_t = x_t^a] + \phi^a + \xi_t^a \tag{23}$$

$$= m_{t-1} + \sum_i \left\{ \lambda(x_t^a)_{it}(x_t^a - E[x_t \mid \Phi_{t-1}, S_t = i])p_{it} \right\} + \phi^a + \varepsilon_t + \xi_t^a \tag{24}$$

$$p_t^b(x_t^b) = E[V_t \mid x_t = x_t^b] - \phi^b + \xi_t^b \tag{25}$$

$$= m_{t-1} + \sum_i \left\{ \lambda(x_t^b)_{it}(x_t^b - E[x_t \mid \Phi_{t-1}, S_t = i])p_{it} \right\} - \phi^b + \varepsilon_t + \xi_t^b \tag{26}$$

where by convention, $x_t^a > 0$ and $x_t^b < 0$. In order to maintain ex-post rationality, the quantities $\lambda(x_t^a)$ and $\lambda(x_t^b)$ represent price impact parameters conditioned on the market maker's quoted depths for the forthcoming trade.

The description of the above process is fully general as it can accommodate the case when trade direction alone is used as the conditioning agent by the market maker to form expectations. In the trade direction case, we simply define the bid and ask 'depths' to equal x_t^b= -1 and x_t^a = 1, respectively.

Other features of the model are also worth noting. The spread attributable to information asymmetry is given by the difference between the quantities that represent the market maker's response to order flow innovation in the bid and ask price:

$$\nabla_{Info} \equiv \sum_i \left\{ \lambda(x_t^a)_{it}(x_t^a - E[x_t \mid \Phi_{t-1}, S_t = i])p_{it} \right\}$$
$$\sum_i \left\{ \lambda(x_t^b)_{it}(x_t^b - E[x_t \mid \Phi_{t-1}, S_t = i])p_{it} \right\} \tag{27}$$

We can see that this quantity is a direct function of the posterior probabilities of the state variable S_t. This suggests that when the market makers' posterior probability of an information event increases, the spread will also adjust to reflect the higher degree of information asymmetry and the increased probability of the presence of informed trading.

The Role of Bounded Rationality in Asset Price formation: The Impact of Net Order Flow and Trade Activity on Prices.

Having specified the mechanism in which the market maker uses aggregated trades over discrete time intervals to infer the underlying state of nature, it is straightforward to see how this will affect pricing. Dependent on which aggregated trade model is being considered, agents within this framework will form the probabilities of S_t not after each trade but after a discrete time period. While these probabilities are updated in a Bayesian fashion, the learning mechanism can be considered to be boundedly rational because the probabilities fail to adjust following the immediate arrival of each trade. A primary motivation for this to occur is that it may be more informative update their beliefs by conditioning on the degree of trade activity that occurs over a given time period. The pricing equations can easily accommodate this type of bounded rational learning process in that the quotes set by the market maker and the assessment of the underlying value are adjusted after each trade, however the probabilities as to which state of nature prevails is adjusted only periodically. This directly impacts upon the market makers expectation of order flow and the error correction mechanism used to form expectations about the security's underlying value. This in turn affects the pricing equations described above.

The model is thus complete. The model of quote setting can now be described as follows. Pricing equations (24), (26), the expectation formation mechanism (22) along with the updating procedure using net order and trade frequency described in Section 1, completely characterize the dynamics of the model.

Given this structure, by observing the realised values for quotes (p_t^a, p_t^b), quoted depths (x_t^a, x_t^b) and actual order flow (x_t), we can apply MCMC techniques to directly generate the posterior densities and perform Bayesian analysis of the structural parameters of the model.

Bayesian Analysis of Bid/Ask Price Formation Models

This section investigates the implementation of MCMC methods to perform Bayesian analysis on the price formation model described. Having obtained the parameters governing the order flow and state variables, we are in a position to investigate the parameters governing the typical pricing equations of (24), (25) and (22) which are re-expressed below:

$$\begin{aligned} p_t^a(x_t^a) &= E[V_t \mid x_t = x_t^a] + \phi^a + \xi_t^a \\ &= m_{t-1} + \sum_i \left\{ \lambda(x_t^a)_{it}(x_t^a - E[x_t \mid \Phi_{t-1}, S_t = i]) p_{it} \right\} + \phi^a + \varepsilon_t + \xi_t^a \\ &= m_{t-1} + g(x_t^a) + \phi^a + u_t^a \end{aligned} \tag{28}$$

$$\begin{aligned} p_t^b(x_t^b) &= E[V_t \mid x_t = x_t^b] - \phi^b + \xi_t^b \\ &= m_{t-1} + \sum_i \left\{ \lambda(x_t^b)_{it}(x_t^b - E[x_t \mid \Phi_{t-1}, S_t = i]) p_{it} \right\} - \phi^b + \varepsilon_t + \xi_t^b , \\ &= m_{t-1} + g(x_t^b) - \phi^b + u_t^b \end{aligned} \tag{29}$$

$$\begin{aligned} m_t &= m_{t-1} + \sum_{i=1}^{3} \left\{ \lambda_{it}(x_t - E[x_t \mid \Phi_{t-1}, s_t = i]) p_{it/t} \right\} + \varepsilon_t \\ &= m_{t-1} + h(x_t) + \varepsilon_t \end{aligned} \tag{30}$$

where $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0,\sigma_\varepsilon^2)$, $u_t^a \sim \mathrm{N}(0,\sigma_{u^a}^2)$, and $u_t^b \sim \mathrm{N}(0,\sigma_{u^b}^2)$ and we define $u_t^b \equiv \varepsilon_t + \xi_t^b$ and $u_t^a \equiv \varepsilon_t + \xi_t^a$.

Let θ be the parameter vector governing the pricing functions (ϕ^b, ϕ^a, λ, $\sigma_\varepsilon^2, \sigma_{u^a}^2, \sigma_{u^b}^2$), where $\lambda = (\lambda_1 \;\; \lambda_2 \;\; \lambda_2)$, and let $Y_n = (y_1,\ldots, y_n)'$ be the vector of observable variables, where $y_t = (p_t^a, p_t^b, x_t)$. Further let $M_t = (m_1,\ldots, m_t)'$ be the vector of unobservable efficient prices. The MCMC strategy augments the parameter space by generating during each sweep of the sampler the latent vector M_t, thereby allowing the full conditional distributions of the parameters to be easily derived. Owing to the linear Gaussian structure adopted in equation (28) to (30), we are able to utilise standard Kalman filtering results to set up an efficient sampling algorithm to sample from the density $f(M_t|\ \theta, Y_n)$. We can then implement the Gibbs Sampler to generate iterates from the joint conditional density $f(\theta, M_t \mid Y_n)$ by drawing samples from $f(\theta \mid M_t, Y_n)$. We first consider the problem of simulating the latent variable vector M_t followed by the derivation of the full conditional posterior densities for the parameter vector which are necessary to implement the Gibbs sampling strategy.

Generating m_t; $t = 1, \ldots, n$ in a state space framework

Following Carter and Kohn (1994), and Chib and Greenberg (1995) we construct an algorithm that samples the vector M_n as a block using the joint conditional distribution, $\Pr(M_n \mid Y_n)$, rather than successively drawing from the set of individual full conditional distributions $\Pr(m_t \mid Y_n, M_{-t})$. Since the stochastic process of the latent variable, m_t, as given in (22) is Markov, and therefore correlated, such blocking will lead to faster convergence to the posterior distribution, and is therefore preferred to the single move sampling.

Given the Gaussian structure of the model, and the timing convention adopted we may reformulate the model so that we can apply Kalman Filtering to generate the state variable. Let

$$y_t \equiv \begin{bmatrix} p_t^a - g(x_t^a) - \phi^a \\ p_t^b - g(x_t^b) + \phi^b \end{bmatrix}, \quad \tilde{x}_t \equiv h(x_t),$$

Then, (28) – (30) can be equivalently expressed in the following state space form:

$$y_{t+1} = H'm_t + w_t \tag{31}$$

$$m_{t+1} = Fm_t + \Gamma x_{t+1} + \varepsilon_t, \tag{32}$$

where

$$F = \Gamma = 1 \qquad H' = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad w_t = \begin{bmatrix} u_t^a \\ u_t^b \end{bmatrix}.$$

Further, given the model specification,

$$E(\varepsilon_t \varepsilon_\tau') = Q\delta_{t\tau}, \; E(w_t w_\tau') = R\delta_{t\tau}, \text{ and } E(w_t \varepsilon_\tau') = S\delta_{t\tau},$$

where,

$Q = \sigma_\varepsilon^2$, $R = \begin{bmatrix} \sigma_{u^a}^2 & 0 \\ 0 & \sigma_{u^b}^2 \end{bmatrix}$, $S = \begin{bmatrix} 0 & 0 \end{bmatrix}$ and δ_{jk} represents Kronecker's delta which equals 1 if $j = k$, and 0 otherwise.

We construct an algorithm which samples the vector M_n as a block using the joint conditional distribution, Pr($M_n | Y_n$), rather than from the set of individual full conditional distributions $\Pr(m_t \mid Y_n, m_{-t})$.

Formally, we generate $M_n \equiv (m_1, m_2, ..., m_n)$ given Y_n, S_n and θ, where $Y_n = (y_1, y_2, ..., y_n)$. Letting $M^t \equiv (m_t, m_{t+1}, ..., m_n)$, $Y^t \equiv (y_t, y_{t+1}, ..., y_n)$ and suppressing θ and S_n for notational convenience, the joint density of M_n can be written as:

$$\Pr(M_n \mid Y_n) = \Pr(m_n \mid Y_n) \prod_{t=1}^{n-1} \Pr(m_t \mid Y_n, M^{t+1}). \tag{33}$$

To generate M_n from Pr($M_n | Y_n$), we first generate m_n from Pr($m_n | Y_n$) and then for t = n-1,…,1 we generate m_t from $\Pr(m_t \mid Y_n, M^{t+1})$ whose density can be derived using Bayes theorem:

$$\begin{aligned} \Pr(m_t \mid Y_n, M^{t+1}) &= \Pr(m_t \mid Y_t, Y^{t+1}, m_{t+1}, M^{t+2}) \\ &\propto \Pr(Y^{t+1}, M^{t+2} \mid Y_t, m_t, m_{t+1}) \Pr(m_{t+1} \mid m_t, Y_t) \Pr(m_t \mid Y_t) \\ &\propto \Pr(m_{t+1} \mid m_t, Y_t) \Pr(m_t \mid Y_t) \end{aligned} \tag{34}$$

since (Y^{t+1}, M^{t+2}) is independent of m_n given (Y_t, m_{t+1}). The density for $m_{t+1} \mid m_t, Y_t$ is ~ N(Fm_t, Q). From the Kalman Filter, the density for $m_t \mid Y_t$ is Gaussian ~ N($\hat{m}_{t|t}, \hat{\Sigma}_{t|t}$), where $\hat{m}_{t|s} = E(m_t \mid Y_s)$ and $\hat{\Sigma}_{t|s} = Var(m_t \mid Y_s)$ for $s \le t \le n$ are obtained by running the following recursions (see Anderson and Moore, 1979):

$$\hat{m}_{t|t} = \hat{m}_{t|t-1} + \hat{\Sigma}_{t|t-1} H (H' \hat{\Sigma}_{t|t-1} H + R)^{-1} (y_{t+1} - H' \hat{m}_{t|t-1}),$$

$$\hat{\Sigma}_{t|t} = \hat{\Sigma}_{t|t-1} - \hat{\Sigma}_{t|t-1} H (H' \hat{\Sigma}_{t|t-1} H + R)^{-1} H' \hat{\Sigma}_{t|t-1},$$

where $\hat{m}_{t+1|t} = F\hat{m}_{t|t} + \Gamma \tilde{x}_{t+1}$, and $\hat{\Sigma}_{t+1|t} = F\hat{\Sigma}_{t|t}F' + Q$.

After the required substitutions into (34), completing the square in m_t leads to the following algorithm to sample {$\tilde{m}_t$} from the joint posterior distribution of states:

1. Run the Kalman filter and compute the values $\hat{\Sigma}_t = \hat{\Sigma}_{t|t} - \tilde{M}_t \hat{\Sigma}_{t+1|t} \tilde{M}_t'$ and $M_t = \hat{\Sigma}_{t|t} \hat{\Sigma}_{t+1|t}^{-1}$.
2. Simulate $\tilde{m}_n$ from N($\hat{m}_{n|n}, \hat{\Sigma}_{n|n}$); then for t = n-1,…,1 generate $\tilde{m}_t$ from N($\hat{m}_t, \hat{\Sigma}_t$), where $\hat{m}_t = \hat{m}_{t|t} + \tilde{M}_t (\tilde{m}_{t+1} - \hat{m}_{t|t})$.

Generating $\theta \mid Y_n, M_n, S_n$

Once M_n (and S_n) have been generated, simulating the individual full conditional distributions for the parameters become relatively straightforward. Given M_n, the complete conditional distribution for the parameters (ϕ^b, ϕ^a, λ, $\sigma_\varepsilon^2, \sigma_{u^a}^2, \sigma_{u^b}^2$) can be easily obtained given the specification of appropriate priors. We can construct the individual full conditional distributions for each parameter in turn, which form the basis of the Gibbs sampling algorithm. The derivations for each of these quantities can be found in the appendix:

$$\lambda \mid Y_n, M_n, S_n, \theta_{-\lambda} \sim \mathrm{N}(\lambda \mid \tilde{\lambda}, \tilde{L}_n^{-1}) \boldsymbol{I}_{\lambda>0} \quad (35)$$

$$\sigma_\varepsilon^2 \mid Y_n, M_n, S_n, \theta_{-\sigma_\varepsilon^2} \sim \mathrm{IG}\left(\frac{\upsilon_1 + n - 1}{2}, \frac{\upsilon_2 + \sum_{t=2}^n (m_t^* - \lambda x_t^*)^2}{2} \right) \quad (36)$$

$$\phi^a \mid Y_n, M_n, S_n, \theta_{-\phi^a} \sim \mathrm{N}(\phi^a \mid \tilde{\phi}^a, \tilde{A}_n^{-1}) \boldsymbol{I}_{\phi^a>0} \quad (37)$$

$$\phi^b \mid Y_n, M_n, S_n, \theta_{-\phi^b} \sim \mathrm{N}(\phi^b \mid \tilde{\phi}^b, \tilde{B}_n^{-1}) \boldsymbol{I}_{\phi^b>0} \quad (38)$$

$$\sigma_{u^a}^2 \mid Y_n, M_n, S_n, \theta_{-\sigma_{u^a}^2} \sim \mathrm{IG}\left(\frac{\upsilon_1^a + n - 1}{2}, \frac{\upsilon_2^a + \sum_{t=2}^n (u_t^a)^2}{2} \right). \quad (39)$$

$$\sigma_{u^b}^2 \mid Y_n, M_n, S_n, \theta_{-\sigma_{u^b}^2} \sim \mathrm{IG}\left(\frac{\upsilon_1^b + n - 1}{2}, \frac{\upsilon_2^b + \sum_{t=2}^n (u_t^b)^2}{2} \right). \quad (40)$$

4. Data and Results

Data Description

The results reported in this study are based on a sample of transactions data taken from the TAQ database, made available by the NYSE. The data comprise of sequenced trade and quotes for the representative stock, Alcoa during the month of September 1994. The stock was selected as it has been the subject of previous microstructure studies (see Manrique and Shephard, 1997).

Following Hasbrouck (1991), Lee and Ready (1991) and Foster and Viswanathan (1993) we adopt a strategy to sign trades whereby transactions at or closest to the bid are treated as seller initiated and hence are negatively signed; trades occurring closest to or at the ask price are considered buyer initiated and hence are positive. Following Easely, Kiefer and O'Hara (1997a,b), when a trade takes place at the midpoint of the quotes, it will be classified depending upon the price movement of the previous trade; if it still can not be determined we move back in the trade record to find the most recent price movement. By doing so we are able to classify every transaction in the sample. All trades during the sample period with the exception of opening trades are thus included for analysis.

In order to implement the model, we must specify the time horizon which the market maker uses to determine the amount of trading activity or trade frequency. The model does not offer any suggestion as to an appropriate choice. So far as that we are studying intraday price dynamics, time intervals shorter than a day would be a logical starting point. As an initial choice 15 minutes was selected as a reasonable horizon. The choice was made based on investigating the average general trading pattern of Alcoa for 1994. The interval seemed to be long enough to preclude the effects of market frictions (such as delays in order submissions and execution, or time lags in the posting of quotes) but short enough to be informative for an intraday analysis. As this choice is arbitrary, the analysis can be carried out using various time horizons to determine the robustness of the model specification. While not reported, preliminary analysis was carried out using a variety of horizons ranging from 1 minute to 1 day. With the exception of the daily time horizon, we obtained qualitatively consistent results to those reported here.

Estimation Results from Net Order Flow Models

We first consider the estimation results obtained by fitting the order flow models to the data using the Gibbs sampling schemes described earlier. An unconditional model was estimated where it was assumed there was no mixture process and compared to the conditional mixture model presented in this study. The MCMC sampler for each model was run for 6000 iterations with the first 1000 discarded.

The posterior means and numerical standard errors for the simulated densities of the models were computed and presented. The numerical standard of the estimates are computed using the heteroskedastic autocorrelation consistent estimator derived by Newey West (1987) to deal with serial correlation in the draws. Other techniques were investigated, including spectral density estimation with Parzen kernels and found the results were similar. We also evaluate the mixture model by comparing the log likelihood and marginal likelihood estimates evaluated at the posterior means using the method developed by Chib (1995). We investigate the efficacy of the Markov mixture representation by comparing them to their respective unconditional specifications. By doing so we can then determine whether the state variable dynamics (S_t) can be adequately described as a Markov mixture (regime switching) process.

Results from Trade Direction Model

Results from the Gibbs sampling scheme for the unconditional and Markov mixture models are presented in Figure 1. Inspection of these plots suggest that the iterates for each parameter have converged to their exact small sample posterior distributions.

Table 1 presents the results of these simulated posteriors for the trade direction model jointly defined by (1) (6) (7) and (9). The results of the Markov mixture model are presented in Panel A with the unconditional model results presented in Panel B. The posterior mean and standard error for the unconditional parameter q (the unconditional probability that incoming trade is a buy) is 0.569 and 1.02 x 10^{-4} respectively. The posterior mean for the average arrival rate of trades per 15 minute period (ϑ) in the unconditional model is 10.68. For the Markov mixture representation the posterior mean of q under a positive information state (S_t=2) is 0.68 , whereas the mean in the negative state (S_t=3) is 0.275. This suggests greater buying pressure in positive states and greater selling pressure during negative states as expected. The posterior mean for ϑ under (S_t=1) (no information event) is 5.30 trades per period which is unambiguously lower than the values reported for S_t=2 and S_t=3 Comparison of the log likelihoods and log marginal likelihoods reported in Panel C of Table 1 overwhelmingly suggests evidence in favour of the mixture model over the unconditional model.

The simulation results for the transition probabilities governing the mixture model are displayed in Figure 2, and Panel A of Table 1. The results suggest that the system remains in state 1 more often than in states 2 and 3. Persistence of the negative state is relatively small as measured by the low transition probability π_{33} = 0.25. Of particular interest are the transition probabilities once an information event has occurred. The results suggest that the probability is high when moving from sate 3 to state 1. However when in state 2 (the positive information state), the probability of moving to state 1 (the no information state) is lower than the probability of moving to sate 3 (the negative information state). These asymmetric results suggest that negative information events are typically followed by periods of calmer trading activity and no news. However when a positive information state prevails, the market is more likely to overshoot and move to a negative state given that the market may have overvalued the security. This notion of over-exuberance during the positive state is also demonstrated by the significantly higher amount of trading activity observed during in these periods; the trade arrival rate in state 2 is roughly 60% higher than the trade arrival rate in state 3.

Figure 3 presents the plots of net order flow and trade activity and the posterior probabilities for each state. By focussing on a smaller sub-period (1 trading week) we can identify the positive information states as those occurring at the beginning and end of a trading day. The probability of no information is low during these periods. No information states are generally found during the middle of the trading day.

Results from Price Formation Model

The price formation model was estimated using one week of trade and quote data during the month of September 1995. A Gibbs sampler was run using 7000 iterations, with the first 2000 iteration discarded. The results of the model estimation using the trade direction are presented in Table 2. The Gibbs sampling posterior densities for these two models are also reproduced in Figures 4 and 5.

Examining the trade direction, we find that the mean estimates for the parameters governing the bid and ask equations produce sensible results; the order processing costs for the ask and bid (ϕ^a and ϕ^b) are approximately 5.95 cents and 5.21 cents per share respectively . These values correspond with the values reported by MRR (1997) with respect to their sample of NYSE stocks. The variances of the bid and ask prices ($\sigma^2_{u^b}$ and $\sigma^2_{u^a}$) are 0.138 and 0.187 cents respectively which are typically small. The variance of the public order flow innovation is 1.85 cents per share and also appears quite small but more significant that the rounding effects implied by $\sigma^2_{u^b}$ and $\sigma^2_{u^a}$. The information asymmetry, or price impact parameters λ_1 λ_2 λ_3 are worthy of note. λ_1 represents the base value that prevails across states and is reported to be 0.038. The marginal price impact parameter values observed under negative and positive states (λ_3 = 0.046 and λ_2 = 0.0017) are strictly positive suggesting greater price sensitivity to trades arriving in information states. The increase in the information asymmetry under these states reflects the greater reliance in the signal content of order flow due to the fact that it indicates the greater possibility of informed trading.

We can observe the price sensitivity of net order flow over time as predicted by the model. The efficient price series obtained using the Kalman filter under each model are presented along with the price sensitivity following each trade in Figure 6, as represented by λ_{CE} the information spread as well as the probability of no information as predicted by each model. What is immediately obvious is that model seems to highlight the increased probability of informed trading at the start and end of each trading day where the probability of no information is low. Investigation of the information spread suggests that the information asymmetries are greatest at the start and end of each trading day as information spreads are higher during those periods. This coincides with the higher probabilities of informed trading observed during these periods. This intraday behaviour is also observed for the certainty equivalent information asymmetry parameter, λ_{CE} .

Conclusions

This study has sought to investigate the way in which information dynamics, order flow and learning combine to influence the formation of prices in financial markets. The notion that market agents condition their learning on a variety of variables has been highlighted by many microstructure studies. In this paper we demonstrate how this can be achieved using a Markov mixture representation of order flow and trade frequency. This representation can in turn be used to describe the dynamics of the latent information process which drives the system. A Bayesian learning mechanism is also introduced to allow market agents to infer the prevailing state of nature given the trading history.

In order to understand how this process affects prices, a simple quote setting and expectation formation model is specified which incorporates the learning mechanism. A major tenet of this model is that agents do not update their beliefs as to the underlying state of nature following the immediate arrival of new information, but after observing the history of trading outcomes over a certain time horizon. This implies that trade activity or the number of trades in a given period is informative. We describe this type of learning as boundedly rational and we test to see if it is an accurate representation pricing behaviour in financial markets.

Upon investigation of these models several conclusions can be drawn. Taken together, the results of the net order flow model suggest that the Markov mixture representation is appropriate in describing the dynamics of information flow. In periods where there is no new private information, the model predicts that trading activity as well as the volatility in order flow will be low. Conversely, during information events, the models predict greater trade activity, and higher volatility, with the sign of net order flow indicating the type of information that is arriving to the market. We find however that the market tends to "over react" following positive information states since negative states are more likely to immediately follow. The converse does not appear to be true with regards to trading behaviour following negative states.

The current approach to modelling order flow and prices has many potential applications in relation to studying market phenomenon over time. In particular, the model can be used to describe how the degree of information asymmetry behaves during certain market events where private information flow is potentially large; such as earnings announcements, M&A announcements and seasoned equity offerings. The model can be extended to other asset classes; in particular bond and derivatives markets. The methodologies developed here can also be applied to existing microstructure models which previously could not be tested directly owing to their latent variable structure. In particular the microstructure model can be extended following Easley and O'Hara (1992) to consider the relative impact of informed and uniformed traders on the price formation process.

In so doing, through the development of more realistic models of price formation, it is hoped that there will be a better understanding of the way in which information and trading behaviour can explain the variation of returns over time and across assets.

References

Albert, J. H., and Chib, S. (1993), "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts", *Journal of Business & Economic Statistics*, **11**, 1-15.

Anderson, B. D. O., and Moore, J. B. (1979), *Optimal Filtering*, Englewood Cliffs, New Jersey: Prentice Hall, Inc.

Carlin, B. P., and Chib, S. (1995), "Bayesian Model Choice via Markov Chain Monte Carlo Method", *Journal of Royal Statistical Society*, Ser.B, **57**, *3*, 473-84.

Carter, C. K., and Kohn, R. (1994), "On Gibbs sampling for state space models", *Biometrika,* **81**, *3*, 541-53.

Casella, G., and George, E. (1992), "Explaining The Gibbs Sampler", *The American Statistician,* **46**, 167-190.

Celeux, G., Hurn, M. and Robert, C.P. (2000) Computational and Inferential Difficulties with mixture posterior distributions, *Journal of the American Statistical Association* , **95**, 957-970.

Chib, S.(1995), "Marginal Likelihood from the Gibbs Output", *Journal of the American Statistical Association,* **90**, 1313-1321.

Chib, S., and Greenberg, E. (1995a), "Understanding the Metropolis-Hastings Algorithm", *The American Statistician,* **49**, *4*, 327-35.

Chib, S., and Greenberg, E. (1995b), "Hierarchical analysis of SUR models with extension to correlated serial errors and time varying parameter models", *Journal of Econometrics,* **68**, 339-60.

Chib, S., and Greenberg, E. (1996), "Markov chain Monte Carlo simulation methods in econometrics", *Econometric Theory*, **12**, 409-431.

Diebolt, J., and Robert, C. P. (1994), “Estimation of finite Mixture Distributions trough Bayesian sampling”, *Journal of Royal Statistical Society*, **56**, *2*, 363-375.

Easley, D., and O’Hara, M. (1987), “Price, trade size, and information in securities markets”, *Journal of Financial Economics*, **19**, 69-90.

Easley, D., and O’Hara, M. (1992), “Time and the Process of Security Price Adjustment”, *Journal of Finance*, **67**, 577-605.

Easley, D., Kiefer, N. M., and O’Hara, M. (1997a), “The information content of the trading process”, *Journal of Empirical Finance*, **4**, 159-86.

Easley, D., Kiefer, N. M., and O’Hara, M. (1997b), “One Day in the Life of a Very Common Stock”, *Review of Financial Studies*, **10**, *3*, 805-35.

Feller, W, (1966), *An introduction to probability theory and its applications*, Vol. 2, 1st ed., New York: Wiley.

Foster, D. F., and Viswanathan, S. (1993), “Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models”, *Journal of Finance*, **48**, *1*, 187-211.

Fruhwirth-Schnatter, S. (2001), “ Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models”, *Journal of the American Statistical Association,* **96**, 194-209.

Gallant, A. R., Rossi, P. E., and Tauchen, G. (1992), “Stock Prices and Volume”, *Review of Financial Studies*, **5**, *2*, 199-242.

Gelfand, A.E. and Smith, A.F.M. (1990), “Sampling Based Approaches to Calculating Marginal Densities”, *Journal of the American Statistical Association*, **85**, 398-409.

Glosten, L. and Harris, L. (1998) “Estimating the components of the bid-ask spread”, *Journal of Financial Economics*, **21**, 123-142.

Glosten, L. R., and Milgrom, P. R. (1985), “Bid/ask and transaction process in a specialist market with heterogeneously informed traders”, *Journal of Financial Economics*, **14**, 71-100.

Gray, S. F. (1996), “Modeling the conditional distribution of interest rates as a regime-switching process”, *Journal of Financial Economics*, **42**, 27-62.

Hamilton, J. D. (1990), “Analysis of Time Series Subject to Changes in Regime”, *Journal of Econometrics*, **45**, 39-70.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton, New Jersey: Princeton University Press.

Han, C., and Carlin, B. P. (2001), “Markov Chain Monte Carlo Methods for Computing Bayes Factors: A Comparative Review”, *Journal of the American Statistical Association*, **96**, *455*, 1122-33.

Hasbrouck, J. (1991), “Measuring the Information Content of Stock Traders”, *Journal of Finance*, **2**, 179-205.

Hasbrouck, J. (1999), “Security bid/ask dynamics with discreetness and clustering: Simple strategies for modelling and estimation”, *Journal of Financial Markets*, **2**, 1-28.

Huang, R. D., and Stoll, H. R. (1997), “The Components of the Bid-Ask Spread: A General Approach”, *Review of the Financial Studies*, **10**, *4*, 995-1034.

Kass, R. E., and Raftery, A. E. (1995), “Bayes Factors”, *Journal of the American Statistical Association*, **90**, *430*, 773-793.

Kavajecz, K. A. (1999), “A Specialist's Quoted Depth and the Limit Order Book”, *Journal of Finance,* **54**, 747-771.

Kyle, A.P. (1985), “Continuous Auctions and Insider Trading”, *Econometrica*, **6**, 1315-1335.

Lee, C. M. C., and Ready, M. J. (1991), “Inferring Trade Direction from Intraday Data”, *Journal of Finance*, **46**, *2*, 733-46.

Lee, C. M. C., Mucklow, B., and Ready, M. J. (1993), “Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis”, *Review of Financial Studies*, **6**, *2*, 345-74.

Madhavan, A. (1992), “Trading Mechanisms in Securities Markets”, *Journal of Finance*, **47**, *2*, 607-41.

Madhavan, A., and Smidt, S. (1991), "A Bayesian model of intraday specialist pricing", *Journal of Financial Economics*, **30**, 99-134.
Madhavan, A., Richardson, M., and Roomans, M. (1997) "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks", *Review of Financial Studies*, **10**, 1035-1064.
Manrique, A., and Shepard, N. (1997), "Likelyhood Analysis of a Discrete Bid/ask Price Model for a Common Stock Quoted on the NYSE", working paper, Nuffield College, Oxford University.
So, M. K. P., Lam, K., and Li, W. K. (1998), "Stochastic Volatility Model With Markov Switching", *Journal of Business & Economic Statistics*, **16**, *2*, 244-253.
Stephens, M. (2000), "Dealing with label switching in mixture models", *Journal of Royal Statistical Society: Ser B*, **62**, *4*, 795-809.
Tanner, M. A. (1996), *Tools for Statistical Inference*, 3rd ed., New York: Springer.

Appendices

Appendix 1: Recursive Equations for Construction of Regime Probabilities

Following Gray (1996), we proceed as follows. First note that if the information set comprises only of the history of trading outcomes, then, $\Phi_{t-1} \equiv X^{t-1} = (X_{t-1},...,X_1)$ and we can equivalently express the regime probabilities as $\Pr(S_t = i \mid \Phi_{t-1}) = \Pr(S_t = i \mid X^{t-1})$. We can then rewrite the probabilities by conditioning on the state at t-1:

$$\begin{aligned}\Pr(S_t = 1 \mid X^{t-1}) &= \sum_{i=1}^{3} \Pr(S_t = 1 | S_{t-1} = i, X^{t-1}) \Pr(S_{t-1} = i | \ X^{t-1}) \\ &= \sum_{i=1}^{3} \Pr(S_t = 1 | S_{t-1} = i) \Pr(S_{t-1} = i | \ X^{t-1}), \qquad \text{(A1.1)}\end{aligned}$$

for $S_t = 1$; similar expressions can be obtained for $\Pr(S_t = 2 \mid X^{t-1})$ and $\Pr(S_t = 3 \mid X^{t-1})$. In vector notation, the vector of probabilities, $\Pr(S_t \mid X^{t-1})$ can be obtained directly from:

$$\Pr(S_t \mid X^{t-1}) = \Pi' \Pr(S_{t-1} \mid X^{t-1}).$$

We are then left to compute $\Pr(s_{t-1} = i \mid X^{t-1})$ for i = 1,2,3. Using Bayes Rule :

$$\begin{aligned}\Pr(S_{t-1} = i \mid X^{t-1}) &= \Pr(S_{t-1} = i \mid x_{t-1}, X^{t-2}) \\ &= \frac{f(X_{t-1} \mid s_{t-1} = i, X^{t-2}) Pr(s_{t-1} = i \mid X^{t-2})}{\sum_{j=1}^{3} f(X_{t-1} \mid s_{t-1} = j, X^{t-2}) Pr(s_{t-1} = j \mid X^{t-2})}, \qquad \text{(A1.2)}\end{aligned}$$

where $f(X_{t-1} \mid S_{t-1} = i, X^{t-2}) = f(X_{t-1} \mid s_{t-1} = i)$, for i = 1,2,3.

We can then form the vector of probabilities, $\Pr(s_{t-1} \mid X^{t-1})$ by computing

$$\Pr(s_{t-1} \mid X^{t-1}) = \left[\frac{f_{t-1} \otimes p_{t-1}}{\iota'(f_{t-1} \otimes p_{t-1})} \right], \qquad \text{(A1.3)}$$

where f_{t-1} represents the vector of likelihoods $f(x_{t-1} \mid s_{t-1} = i, X^{t-2})$, for i= 1,2,3 at t-1, and p_{t-1} is the vector of probabilities $\Pr(S_{t-1} = i \mid X^{t-2})$. By substituting (A1.2) into (A1.3) then yields the desired expression to generate the regime probabilities, $p_{it} = \Pr(S_t = i | X^{t-1})$ for i = 1,2,3, as given by equation (3.4)

Appendix 2: Deriving the Conditional Distribution for Net Order Flow in A Trade Direction Model

In order to derive the distribution for this variable, we can view Q_t = $\sum_{\tau=1}^{n_t} x_{\tau,t}$ as a simple binomial random walk process where the order flow variable $\{x_{\tau,t}\}$ is a sequence of mutually independent identically distributed random variables. The possible values of the random variable, Q_t, after n_t trials are $k = 0, \pm 1 \ldots \pm n_t$. In order to derive the probability that $Q_t = k$, we can rewrite $x_{\tau,t}$ as $2\tilde{x}_{\tau,t} - 1$ where $\tilde{x}_{\tau,t}$ is a Bernoulli random variable. As a result we know that $\sum_{\tau=1}^{n_t} \tilde{x}_{\tau,t}$ will be Binomial:

$$p(\tilde{Q}_t \equiv \sum_{\tau=1}^{n_t} \tilde{x}_{\tau,t} \mid n_t, s_t = i) = \binom{n_t}{\tilde{Q}_t} p_i^{\tilde{Q}_t} q_i^{n_t - \tilde{Q}_t} . \tag{A2.1}$$

Since $Q_t = \sum_{\tau=1}^{n_t} x_{\tau,t} = \sum_{\tau=1}^{n_t}(2\tilde{x}_{\tau,t} - 1) = 2\tilde{Q}_t - n_t$, then

$$\begin{aligned} p(Q_t \mid n_t, s_t = i) &= p(\tilde{Q}_t = \frac{1}{2}(Q_t + n_t) \mid n_t, q_i) \\ &= \binom{n_t}{\frac{1}{2}(n_t + Q_t)} p_i^{\frac{1}{2}[n_t + Q_t]} q_i^{\frac{1}{2}[n_t - Q_t]} . \\ &= \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{Q_t + v_t} q_i^{v_t} \end{aligned} \tag{A2.2}$$

where we define $n_t - Q_t = 2v_t$, such that the probability vanishes for odd k when n_t is even and for even k when n_t is odd. Hence

$$p(Q_t, n_t \mid s_t = i) = \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{Q_t + v_t} q_i^{v_t} . \frac{e^{-\vartheta_i} \vartheta_i^{n_t}}{n_t!} \tag{A2.3}$$

It can be shown (see Feller, 1966, pg 59) that the marginal pdf for Q_t without conditioning on n_t is given by

$$\begin{aligned} p(Q_t \mid s_t = i) &= \sum_{n_t=0}^{\infty} \frac{e^{-\vartheta_i} \vartheta_i^{n_t}}{n_t!} \binom{Q_t + 2v_t}{Q_t + v_t} p_i^{Q_t + v_t} q_i^{v_t} \\ &= \sqrt{(p_i / q_i)^{Q_t}}\, e^{-\vartheta_i} I_{Q_t}(2\sqrt{p_i q_i}\, \vartheta_i) , \end{aligned} \tag{A2.4}$$

where we define the function I_ρ for all real x as the modified Bessel function:

$$I_\rho(x) = \sum_{k=0}^{\infty} \frac{1}{k!\,\Gamma(\mathrm{k} + \rho + 1)} \left(\frac{x}{2}\right)^{2k+\rho} . \tag{A2.5}$$

We note that under the special case that $n_t = 0$ (which implies $Q_t = 0$), both the joint and marginal densities are well defined, thus allowing an analysis of cases of when no trade intervals occur.

Appendix 3: Full Conditional Distributions for the parameters and Latent Variables in Net Order Flow Model

A3.1 Block Sampling Algorithm For The Generation of Discrete State Vector, *S*

Following Carter and Kohn (1994), the algorithm first generates S_n from $\Pr(S_n \mid X_n)$ and then for $t = n$-1,…,1 successively generates S_t from $\Pr(S_t \mid X_t , S_{t+1})$ which from Bayes theorem, for $t = n$-1,…,1 can be computed from :

$$\Pr(S_t \mid X_t , S_{t+1}) = \frac{\Pr(S_{t+1} \mid S_t)\Pr(S_t \mid X_t)}{\Pr(S_{t+1} \mid X_t)}. \qquad \text{(A3.1)}$$

In order to compute this quantity, $\Pr(S_t \mid X_t)$ must be generated. The discrete filter to evaluate $\Pr(S_t \mid X_t)$ can then be described as a set of recursive equations used generate samples from $\Pr(S^n \mid X_n)$. Noting that $\Pr(S_t \mid X^t) \propto \Pr(X_t \mid X^{t-1}, S_t)\Pr(S_t \mid X_{t-1})$, the discrete filter to obtain $\Pr(S_t \mid X_{t-1})$ and $\Pr(S_t \mid X_t)$ can be obtained as follows for $t = 1,\ldots,$n :

1 Compute $\Pr(S_t \mid X_{t-1}) = \sum_j \Pr(S_t \mid S_{t-1} = j)\Pr(S_{t-1} = j \mid X_{t-1})$.

2. Set $p^*(S_t \mid X_t) = \Pr(X_t \mid X_{t-1}, S_t)\Pr(S_t \mid X_{t-1})$.

3. Obtain $\Pr(S_t \mid X_t)$ using

$$\Pr(S_t \mid X_t) = p^*(S_t \mid X_t) \Big/ \sum_j p^*(S_t = j \mid X_t).$$

These results can then be fed into the Bayes expression (A3.1) to sample S^n from $\Pr(S^n \mid X_n)$.

A3.2 Full Conditional Distribution of Π_i

From Bayes Rule we can express the posterior density of Π_i as

$$\Pr(\Pi_i \mid S^n) \propto \Pr(S^n \mid \Pi_i)\Pr(\Pi_i).$$

Given that S_t evolves according to a first order Markov process, the joint likelihood for S^n is Dirichlet. By adopting conjugate priors for $\Pr(\Pi_i)$, the posterior denisty too will be a dirichlet, and so that parameters for Π_i can be jointly sampled from the following Dirichlet distribution :

$$\Pr(S^n \mid \Pi_i) \propto p_{i1}^{n_{i1}} p_{i2}^{n_{i2}} \ldots p_{ik}^{n_{ik}} = \prod_{j=1}^{k-1} p_{ij}^{n_{ij}} .(1 - \sum_{j=1}^{k-1} p_{ij})^{n_{ik}} \qquad \text{(A3.2)}$$

where n_{ij} represents the number of transitions from state i to state j; $n_{ij} = \sum_{t=2}^{n} \boldsymbol{I}(s_{t-1} = i)\boldsymbol{I}(s_t = j)$, and k = 3 is the number of states. By adopting conjugate priors for $\Pr(\Pi_i) \propto p_{i1}^{u_{i1}} .p_{i2}^{u_{i2}} ...p_{iK}^{u_{iK}}$, where u_{ij} are the hyperparameters of the Dirichlet prior, the posterior distribution is given by

$$\Pr(\Pi_i \mid S_n) \propto p_{i1}^{n_{i1}+u_{i1}} .p_{i2}^{n_{i2}+u_{i2}} ...p_{ik}^{n_{iK}+u_{ik}}. \qquad \text{(A3.3)}$$

Hence the posterior is a Dirichlet distribution $\mathrm{Dir}(d_{i1}, d_{i2}, \ldots, d_{ik})$ where $d_{ij} = n_{ij} + u_{ij}$.

A3.3 Full Conditional Distribution of q_i

From Bayes theorem the posterior for q_i can be expressed as :

$$p(p_i \mid X^T, S_T, \vartheta_i) \propto p(X^T \mid S_T, p_i, \vartheta_i)p(p_i).$$

In order to obtain the posterior density, an appropriate prior must be chosen for p_i. A useful choice for the vector would be the conjugate beta prior given that each element of p_i represents a probability bounded between zero and one and that the vector must sum to one; i.e. p_i is distributed as Beta (b_i , s_i) with hyperparameters, b_i , s_i . Thus from the joint sampling density and the prior we obtain the posterior density;

$$\text{Beta}(b_i + \textstyle\sum_{t=1}^{T}(Q_t + v_t)\, , s_i + \sum_{t=1}^{T} v_t). \qquad \text{(A3.4)}$$

A3.4 Full Conditional Distribution of ϑ_i

From Bayes theorem the posterior for ϑ_i can be given by :

$$p(\vartheta_i \mid X^T, S_T, q_i) \propto p(X^T \mid S_T, q_i)\, p(\vartheta_i).$$

In order to obtain the posterior density, an appropriate prior must be chosen for ϑ_i. A useful choice would be a conjugate gamma prior: $p(\vartheta_i) \propto \vartheta_i^{c_i-1} e^{-d_i\vartheta_i}$. Thus from the joint sampling density and the prior we obtain the posterior density as:

$$\begin{aligned} p(\vartheta_i \mid X^T, S_T, q_i) &\propto \vartheta_i^{\sum_{t=1}^{T} n_t} e^{-T\vartheta_i}\ \vartheta_i^{c_i-1} e^{-d_i\vartheta_i} \\ &\propto \vartheta_i^{\sum_{t=1}^{T} n_t + c_i - 1} e^{-(T+d_i)\vartheta_i}. \end{aligned} \qquad \text{(A3.5)}$$

Therefore the posterior density is Gamma ($\sum_{t=1}^{T} n_t + c_i, T + d_i$).

Appendix 4: Full Conditional Distributions for the parameters and Latent Variables in Price Formation Models

A.4.1 Full Conditional Distribution of σ_ε^2 |Y_n, M_n, S_n, $\theta_{-\sigma_\varepsilon^2}$

Consider the state equation

$$m_t = m_{t-1} + \sum_{i=1}^{3} \left\{ \lambda_{it}(x_t - E[x_t \mid \Phi_{t-1}, s_t = i]) p_{it/t} \right\} + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0, \sigma_\varepsilon^2).$$

By conditioning on the latent variable, which is generated via the Kalman Filter, we can express the posterior distribution of σ_ε^2 as

$$\begin{aligned} \Pr(\sigma_\varepsilon^2 \mid Y_n, S_n, M_n, \theta_{-\sigma_\varepsilon^2}) &= \Pr(\sigma_\varepsilon^2 \mid X_n, M_n, \theta_{-\sigma_\varepsilon^2}), \text{ where } X_n \equiv (x_1, x_2, \dots x_n)' \\ &\propto \Pr(M_n \mid X_n, \theta)\Pr(X_n \mid \theta)\Pr(\theta) \\ &\propto \Pr(M_n \mid X_n, \theta)\Pr(\sigma_\varepsilon^2), \end{aligned}$$

since X_n is independent of σ_ε^2. Let the prior distribution of $\sigma_\varepsilon^2 \sim \mathrm{IG}(\upsilon_1/2, \upsilon_2/2)$.

Since $\Pr(M_n \mid X_n, \theta)$ can be expressed as

$\Pr(m_1,\dots,m_n \mid X_n, \theta) = \Pr(m_1 \mid X_n, \theta) \prod_{t=2}^{n} \Pr(m_t \mid m_{t-1}, X_n, \theta)$, by conditioning on m_1, and defining $m_t^* \equiv m_t - m_{t-1}$ and $x_t^* \equiv x_t - \mathrm{E}(x_t \mid \Phi_{t-1})$,

$$\Pr(M_n \mid X_n, \theta) \propto (\sigma_\varepsilon^2)^{-(n-1)/2} \exp\{-1/2\sigma_\varepsilon^2 \textstyle\sum_{t=2}^{n} (m_t^* - \lambda x_t^*)^2\}.$$

The complete distribution of σ_ε^2 can then be given by:

$$\sigma_\varepsilon^2 \mid Y_n, M_n, S_n, \theta_{-\sigma_\varepsilon^2} \sim \mathrm{IG}\left(\frac{\upsilon_1 + n - 1}{2}, \frac{\upsilon_2 + \sum_{t=2}^{n} (m_t^* - \lambda x_t^*)^2}{2} \right). \qquad \text{(A4.1)}$$

A.4.2 Full Conditionals of $\sigma_{u^a}^2$ |Y_n, M_n, S_n, $\theta_{-\sigma_{u^a}^2}$ and $\sigma_{u^b}^2$ |Y_n, M_n, S_n, $\theta_{-\sigma_{u^b}^2}$

We first consider in detail, $\sigma_{\xi^a}^2$ |Y_n, M_n, S_n, $\theta_{-\sigma_{u^a}^2}$. Along similar lines to σ_ε^2, we can express the posterior distribution of $\sigma_{u^a}^2$ as

$$\begin{aligned} \Pr(\sigma_{u^a}^2 \mid Y_n, S_n, M_n, \theta_{-\sigma^2}) &= \Pr(\sigma_{u^a}^2 \mid Y_n^1, M_n, \theta_{-\sigma_{u^a}^2}), \text{ where } Y_n^1 \equiv (p_1^a, p_2^a, \dots, p_n^a) \\ &\propto \Pr(Y_n^1 \mid M_n, \theta)\Pr(M_n \mid \theta)\Pr(\theta) \\ &\propto \Pr(Y_n^1 \mid M_n, \theta)\Pr(\sigma_{u^a}^2), \end{aligned}$$

since M_n is independent of $\sigma_{u^a}^2$.

Letting the prior distribution of $\sigma_{u^a}^2 \sim \mathrm{IG}(\upsilon_1^a/2, \upsilon_2^a/2)$, $\Pr(Y_n^1 \mid M_n, \theta)$ can be expressed as

$$\Pr(Y_n^1 \mid M_n, \theta) = \Pr(p_1^a, \dots, p_n^a \mid M_n, \theta)$$

$$= \Pr(p_1^a \mid m_0, \theta) \prod\nolimits_{t=2}^{n} \Pr(p_t^a \mid m_{t-1}, \theta) .$$

Conditioning on p_1^a, and letting

$$u_t^a = p_t^a - m_{t-1} - \sum_i \left\{ \lambda_{it} (x_t^a - E[x_t \mid \Phi_{t-1}, S_t = i]) p_{it} \right\} - \phi^a ,$$

$$\Pr(Y_n^1 \mid M_n, \theta) \propto (\sigma_{u^a}^2)^{-(n-1)/2} \exp\{-\frac{1}{2\sigma_{u^a}^2} \sum\nolimits_{t=2}^{n} (u_t^a)^2 \} .$$

We can then show that the complete distribution of $\sigma_{u^a}^2$ is given by:

$$\sigma_{u^a}^2 \ |Y_n, M_n, S_n, \ \theta_{-\sigma_{\xi^a}^2} \sim \ \mathrm{IG}\left(\frac{\upsilon_1^a + n - 1}{2}, \frac{\upsilon_2^a + \sum\nolimits_{t=2}^{n} (\xi_t^a)^2}{2} \right). \qquad \text{(A4.2)}$$

Similarly for $\sigma_{u^b}^2$ $|Y_n, M_n, S_n, \ \theta_{-\sigma_{u^b}^2}$, assuming a prior for $\sigma_{u^b}^2$ ~ IG($\upsilon_1^b/2, \upsilon_2^b/2$), we can obtain the conditional distribution as

$$\sigma_{u^b}^2 \ |Y_n, M_n, S_n, \ \theta_{-\sigma_{u^b}^2} \sim \ \mathrm{IG}\left(\frac{\upsilon_1^b + n - 1}{2}, \frac{\upsilon_2^b + \sum\nolimits_{t=2}^{n} (u_t^b)^2}{2} \right). \qquad \text{(A4.3)}$$

A.4.3 Full Conditional Distribution of λ $|Y_n, M_n, S_n, \ \theta_{-\lambda}$

We first define $\lambda = (\lambda_1 \quad \lambda_2 \quad \lambda_3)$. From the expectation formation equation:

$$m_t = m_{t-1} + \sum_{i=1}^{3} \left\{ \lambda_{it} (x_t - E[x_t \mid \Phi_{t-1}, s_t = i]) p_{it/t} \right\} + \varepsilon_t , \qquad \text{(A4.4)}$$

where $$\lambda_{it} = \begin{cases} \lambda_1 & \text{if} \quad S_t = 1 \\ \lambda_1 + \lambda_2 g_2(x_t) & \text{if} \quad S_t = 2 \\ \lambda_1 + \lambda_3 g_3(x_t) & \text{if} \quad S_t = 3 \end{cases} \qquad \text{(A4.5)}$$

The functions $g_i(.)$ will depend upon whether we are dealing with a trade indicator variable for x_t or a signed trade size variable. On substitution of (A4.4) into this (A4.5) and simplifying yields

$$m_t = m_{t-1} + \lambda_1 z_{1t} + \lambda_2 z_{2t} + \lambda_2 z_{2t} + \varepsilon_t , \qquad \text{(A4.6)}$$

where $z_{1t} = \sum_{i=1}^{3} \hat{x}_{it}$, $z_{2t} = g_2(x_t).\hat{x}_{2t}$, $z_{3t} = g_3(x_t).\hat{x}_{3t}$, and

$$\hat{x}_{it} \equiv x_t - E[x_t \mid \Phi_{t-1}, s_t = i]) p_{it/t} .$$

We can now express the posterior distribution for λ as

$$\Pr(\lambda \mid Y_n, S_n, M_n, \theta_\lambda) \propto \Pr(M_n \mid X_n, \theta) \Pr(X_n \mid \theta) \Pr(\theta)$$
$$\propto \Pr(M_n \mid X_n, \theta) \Pr(\lambda) .$$

Assuming the truncated prior $\lambda \sim \mathrm{N}(\lambda \mid \lambda_0, L_0^{-1}) \boldsymbol{I}_{\lambda>0}$, where $\boldsymbol{I}$ is the indicator function on [λ>0], it can be shown after conditioning on m_1, that the full conditional of λ can be expressed as (see Chib and Greenberg, 1996) :

$$\lambda \ |Y_n, M_n, S_n, \ \theta_{-\lambda} \sim \mathrm{N}(\lambda \mid \tilde{\lambda}, \tilde{L}_n^{-1}) \boldsymbol{I}_{\lambda>0} \qquad \text{(A4.7)}$$

where $\tilde{L}_n = (L_0 + \sigma_\varepsilon^{-2} Z'Z)$, $\tilde{\lambda} = \tilde{L}_n^{-1}(L_0 \tilde{\lambda} + \sigma_\varepsilon^{-2} Z' m^*)$, $Z_t \equiv (z_{1t} \ z_{2t} \ z_{3t})$, $Z \equiv (Z_2^{'}, \ldots Z_n^{'})$, and $m^* \equiv (m_2^*, \ldots m_n^*)'$.

A.4.4 Full Conditionals of ϕ^a $|Y_n, M_n, S_n, \ \theta_{-\phi^a}$, and ϕ^b $|Y_n, M_n, S_n, \ \theta_{-\phi^b}$

To obtain the full conditional for ϕ^a and ϕ^b, we focus on pricing equations:

$$p_t^a(x_t^a) = m_{t-1} + \sum_i \left\{ \lambda_{it}(x_t^a - E[x_t \mid \Phi_{t-1}, S_t = i]) p_{it} \right\} + \phi^a + u_t^a$$
$$p_t^b(x_t^b) = m_{t-1} + \sum_i \left\{ \lambda_{it}(x_t^b - E[x_t \mid \Phi_{t-1}, S_t = i]) p_{it} \right\} - \phi^b + u_t^b .$$

Let the conjugate priors for ϕ^a and ϕ^b be given by truncated normals:

$$\phi^a \sim N(\phi^a \mid \phi_0^a, A_0^{-1}) \boldsymbol{I}_{\phi^a > 0} \text{ and } \phi^b \sim N(\phi^b \mid \phi_0^b, B_0^{-1}) \boldsymbol{I}_{\phi^b > 0} .$$

We first consider the full conditional distribution for ϕ^a. A similar expression can then be obtained for ϕ^b. Define $\overline{p}_t^a = p_t^a - m_{t-1} - \sum_i \left\{ \lambda_{it}(x_t^a - E[x_t \mid \Phi_{t-1}, S_t = i]) p_{it} \right\}$. It can then be shown after conditioning on $\overline{p}_1^a$, that the full conditional of ϕ^a is expressed as (see Chib and Greenberg, 1996):

$$\phi^a \mid Y_n, M_n, S_n, \theta_{-\phi^a} \sim \mathrm{N}(\phi^a \mid \widetilde{\phi}^a, \widetilde{A}_n^{-1}) \boldsymbol{I}_{\phi^a > \mathbf{0}} \quad \text{(A4.8)}$$

where $\widetilde{A}_n = (A_0 + \sigma_{u^a}^{-2} n)$, $\widetilde{\phi}^a = \widetilde{A}_n^{-1}(A_0 \phi_0^a + \sigma_{u^a}^{-2} \sum_{t=2}^{n} \overline{p}_t^a)$ and similarly

$$\phi^b \mid Y_n, M_n, S_n, \theta_{-\phi^b} \sim \mathrm{N}(\phi^b \mid \widetilde{\phi}^b, \widetilde{B}_n^{-1}) \boldsymbol{I}_{\phi^b > \mathbf{0}} \quad \text{(A4.9)}$$

where $\widetilde{B}_n = (B_0 + \sigma_{u^a}^{-2} n)$, $\widetilde{\phi}^b = \widetilde{B}_n^{-1}(B_0 \phi_0^b + \sigma_{u^b}^{-2} \sum_{t=2}^{n} \overline{p}_t^b)$.

Table 1. Summary of Posterior Parameter Distributions for Trade Direction Based Model of Net Order Flow

This table reports summary statistics for the simulated posterior distributions of the model of net order flow where direction is used as the basis for the order flow variable. The probability that an incoming trade is buyer initiated is given by q, and the mean arrival rate of trades over a 15 minute trading interval is given by ϑ. For the Markov mixture model, these parameters are subscripted to reflect each regime. Transition probabilities, π_{ij}, for the Markov mixture model are also presented. The posteriors were constructed using a Gibbs sampler over 6000 iterations, with the first 1000 discarded. Panel A reports the posterior mean, posterior 95% intervals, and numerical standard error for the posterior mean estimate under the 3 regime Markov Mixture Model. Panel B reports the estimation results for the unconditional model. Panel C compares the two specifications by reporting the log of the marginal likelihood and log likelihood evaluated at the posterior means for each model. Numerical Standard Errors are computed using a heteroskedastic autocorrelation consistent estimator (Newey-West).

Panel A (Mixture Model)

Parameter	Posterior Mean	Posterior 95% Interval		Numerical Standard Error
q_1	0.642	0.5839	0.7103	1.36×10^{-3}
q_2	0.679	0.6511	0.7022	4.83×10^{-4}
q_3	0.275	0.2292	0.3220	7.42×10^{-4}
ϑ_1	5.300	4.6989	6.1304	0.016
ϑ_2	17.614	16.1438	19.8605	0.045
ϑ_3	11.322	9.8498	12.7600	0.031
π_{11}	0.644	0.5454	0.7285	1.19×10^{-3}
π_{12}	0.169	0.1048	0.2458	1.01×10^{-3}
π_{13}	0.187	0.1034	0.3021	1.86×10^{-3}
π_{21}	0.177	0.1064	0.2604	7.98×10^{-4}
π_{22}	0.519	0.4258	0.6130	8.52×10^{-4}
π_{23}	0.303	0.2087	0.4045	9.13×10^{-4}
π_{31}	0.405	0.2860	0.5351	1.84×10^{-3}
π_{32}	0.260	0.1878	0.4751	2.90×10^{-3}
π_{33}	0.260	0.1400	0.3843	1.99×10^{-3}

Panel B (Unconditional model)

Parameter	Posterior Mean	Posterior 95% Interval		Numerical Standard Error
q	0.569	0.5559	0.5824	1.02×10^{-4}
ϑ	10.680	10.3918	10.9750	1.73×10^{-3}

Panel C (Comparison of Mixture and Unconditional Models)

	Mixture	Unconditional Model
Log Marginal Likelihood	-5063	-5840
Log Likelihood	-4712	-5730

Table 2. Summary of Posterior Parameter Distributions for Quote Setting Model with Trade Direction as basis for Net Order Flow

This table reports summary statistics for the simulated posterior distributions of the quote setting model of net order flow where trade direction is used as the basis for the order flow variable. The posteriors were constructed using a Gibbs sampler over 3000 iterations, with the first 1000 discarded. The log likelihood evaluated at the posterior means of the parameters is also reported. Numerical Standard Errors are computed using a heteroskedastic autocorrelation consistent estimator (Newey-West).

Parameter	Posterior Mean	Posterior 90% Interval		Numerical Standard Error
Bid and Ask Equation Parameters				
ϕ^a	0.05946	0.05302	0.06585	0.00024
ϕ^b	-0.05208	-0.05814	-0.04621	0.00024
$\sigma^2_{u^a}$	0.00187	0.00163	0.00212	7.560×10^{-6}
$\sigma^2_{u^b}$	0.00138	0.00118	0.00159	6.713×10^{-6}
Expectation Formation Parameters				
σ^2_{ε}	0.01847	0.01588	0.02108	2.833×10^{-5}
λ_1	0.03832	0.03125	0.04565	7.317×10^{-5}
λ_2	0.04550	0.03416	0.05689	9.947×10^{-5}
λ_3	0.00169	0.00153	0.00187	2.526×10^{-6}
Marginal Log Likelihood		4773.33		

Figure 1. Simulated Parameters for the Unconditional Model of Net Order Flow, $Q_t = \sum_{\tau=1}^{n_t} x_{\tau,t}$ (Informative Variable: $x_{\tau t}$ = Trade Direction).

This figure reports the simulated posterior distributions for the parameters of the unconditional net order flow model. The results are based on a Gibbs sampler run of 6000 iterations with the first 1000 discarded. For each parameter we report the iterates and the histograms of the marginal distributions resulting from the Markov Chain. Panel A reports the posterior density for the unconditional probability of buyer initiated trade q. Panel B reports the posterior density for the unconditional Poisson parameter, ϑ, that governs the trade activity variable, n_t.

Iterates from Gibbs Run *Histogram*

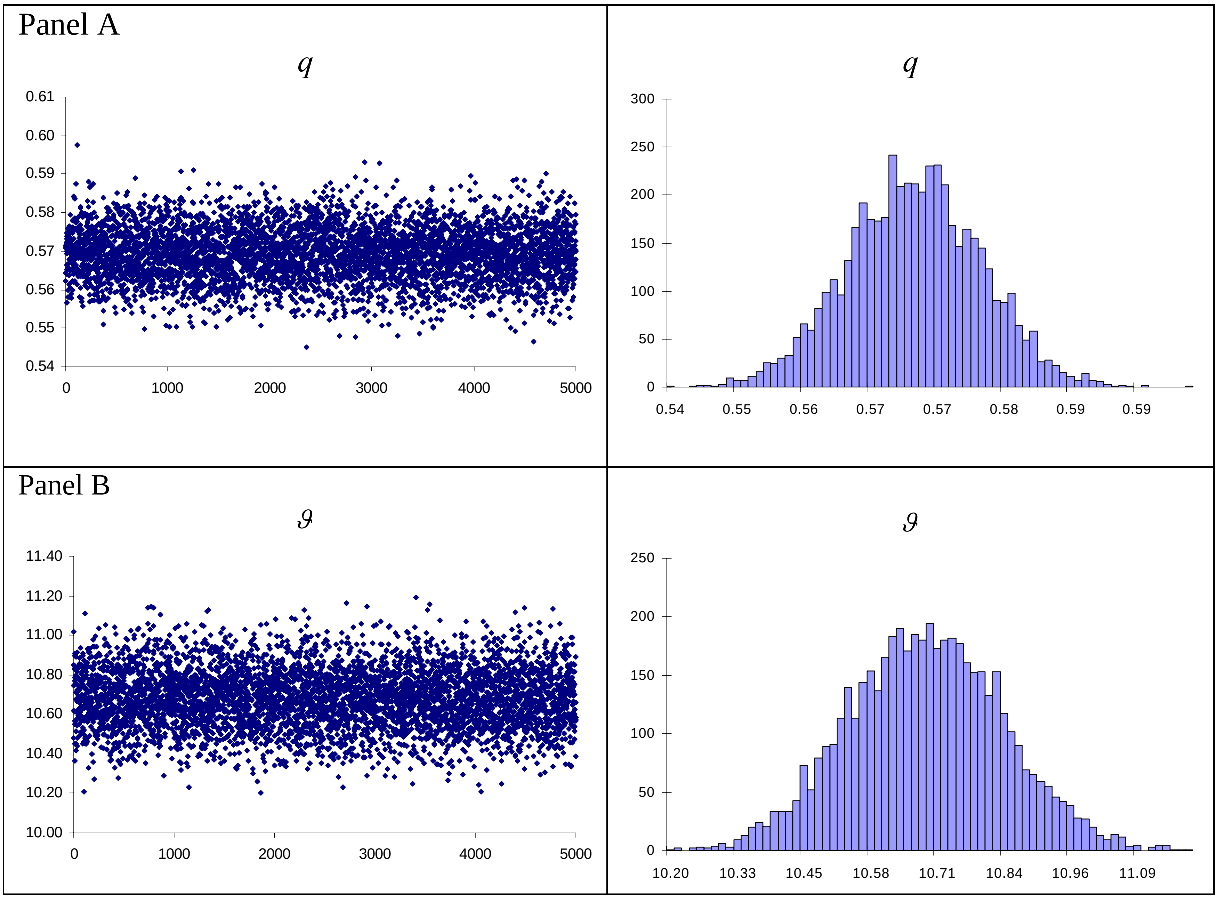


Figure 2. Simulated Parameters for the Markov Mixture Model of Net Order Flow, $Q_t = \sum_{\tau=1}^{n_t} x_{\tau,t}$ (Informative Variable: $x_{\tau t}$ = Trade Direction).

This figure reports the simulated posterior distributions for the parameters of the Markov mixture model of net order flow under each regime $S_t = i$; where i can take on the values: 1 (no information event), 2 (positive information event) , and 3 (negative information event). The results are based on a Gibbs sampler run of 6000 iterations with the first 1000 discarded. For each parameter, we report the iterates and the histograms of the marginal distributions resulting from the Markov Chain. Panel A reports the posterior densities for the conditional probability of buyer initiated trade q_i Panel B reports the posterior densities for the conditional Poisson parameters, ϑ_i, that govern trade activity variable, n_t.

Iterates from Gibbs Run *Histogram*

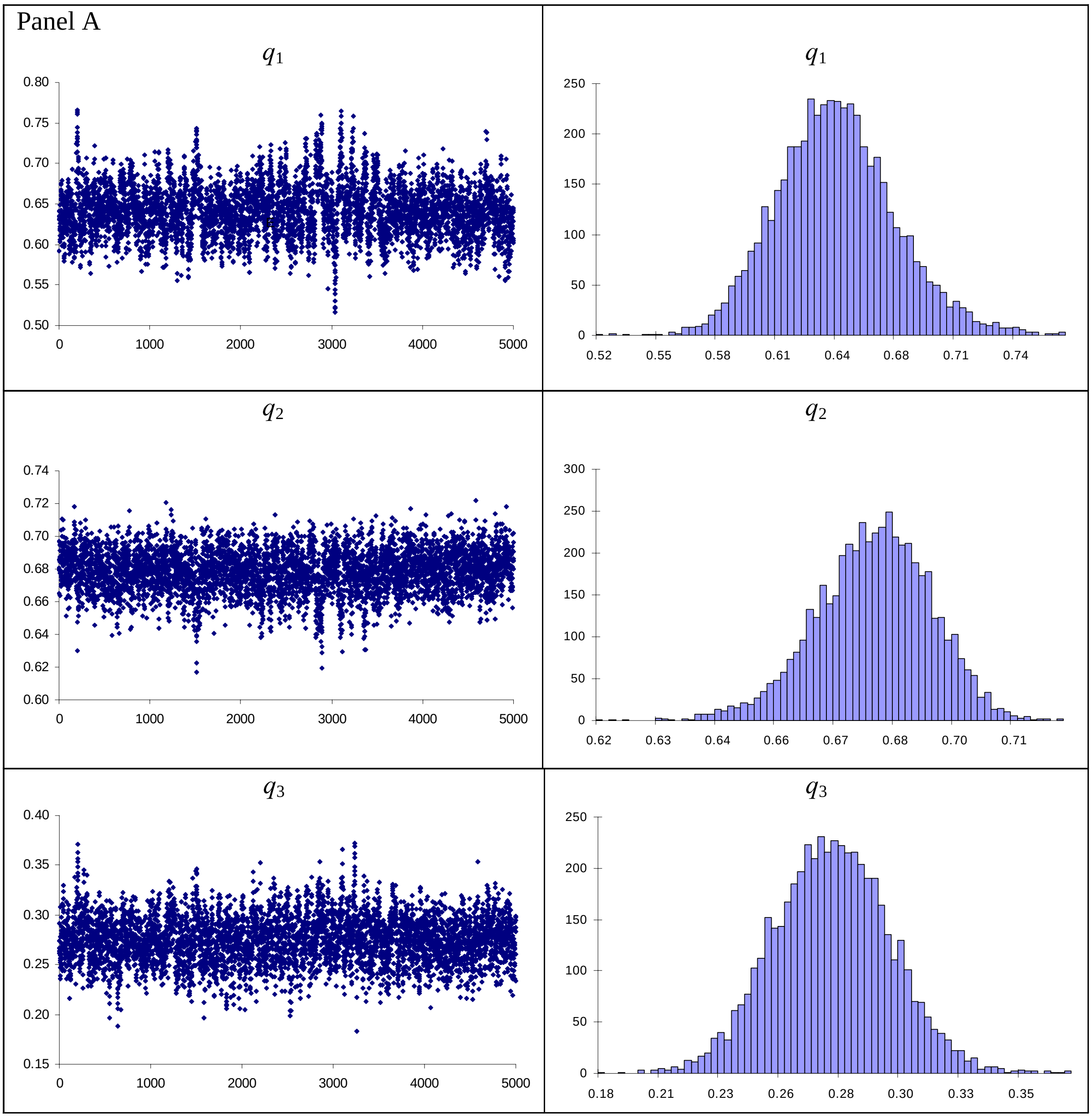


Figure 2 continued.

Iterates from Gibbs Run *Histogram*

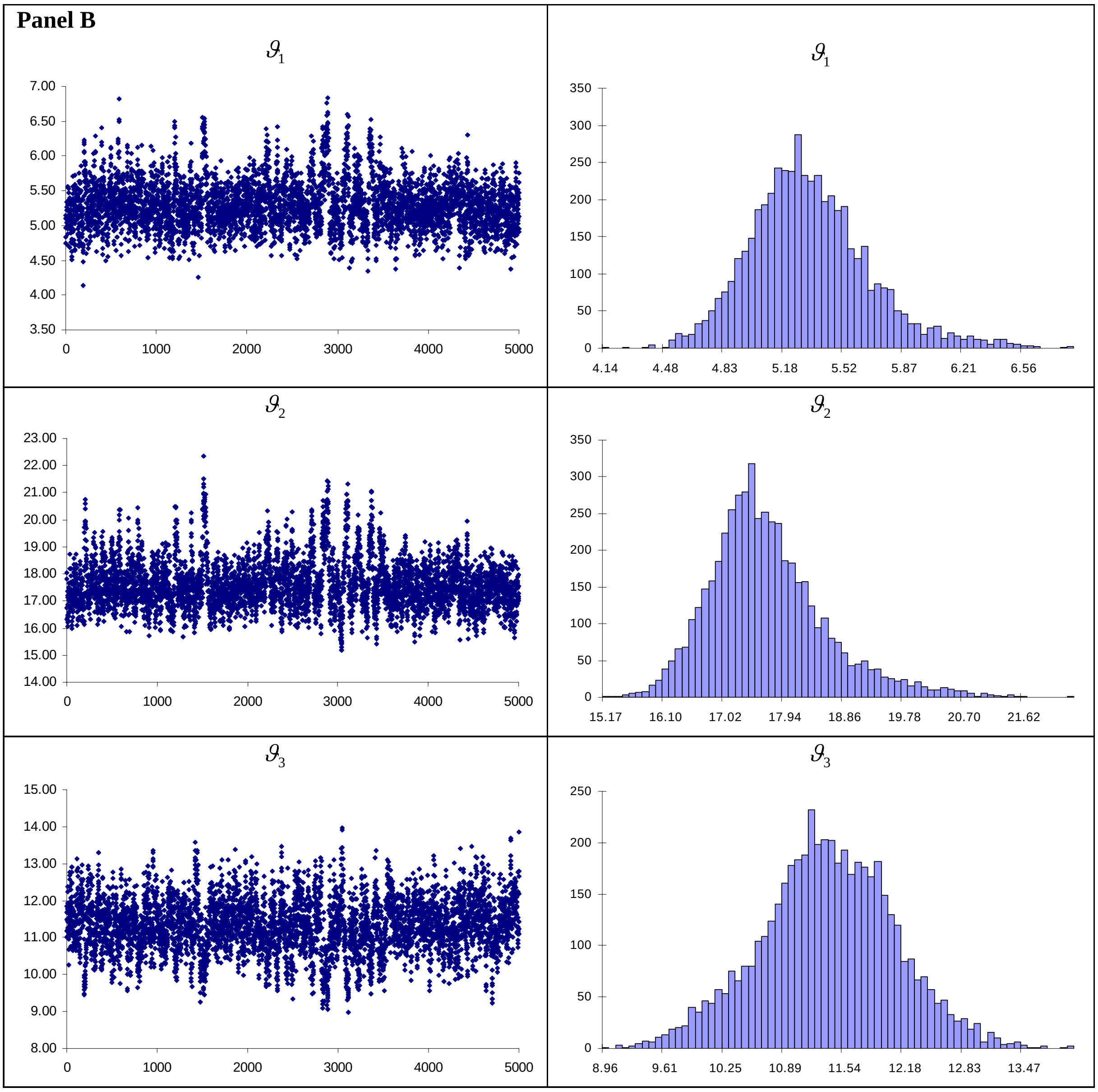


Figure 3. Posterior densities of the transition probabilities for trade direction based model of Net Order Flow.

The figure is based on a Gibbs sampler run of 6000 iterations with the first 1000 discarded. The histograms presented are the resulting marginal distributions for each π_{ij} where $\pi_{ij} = \Pr(S_t = j \mid S_{t-1} = i)$.

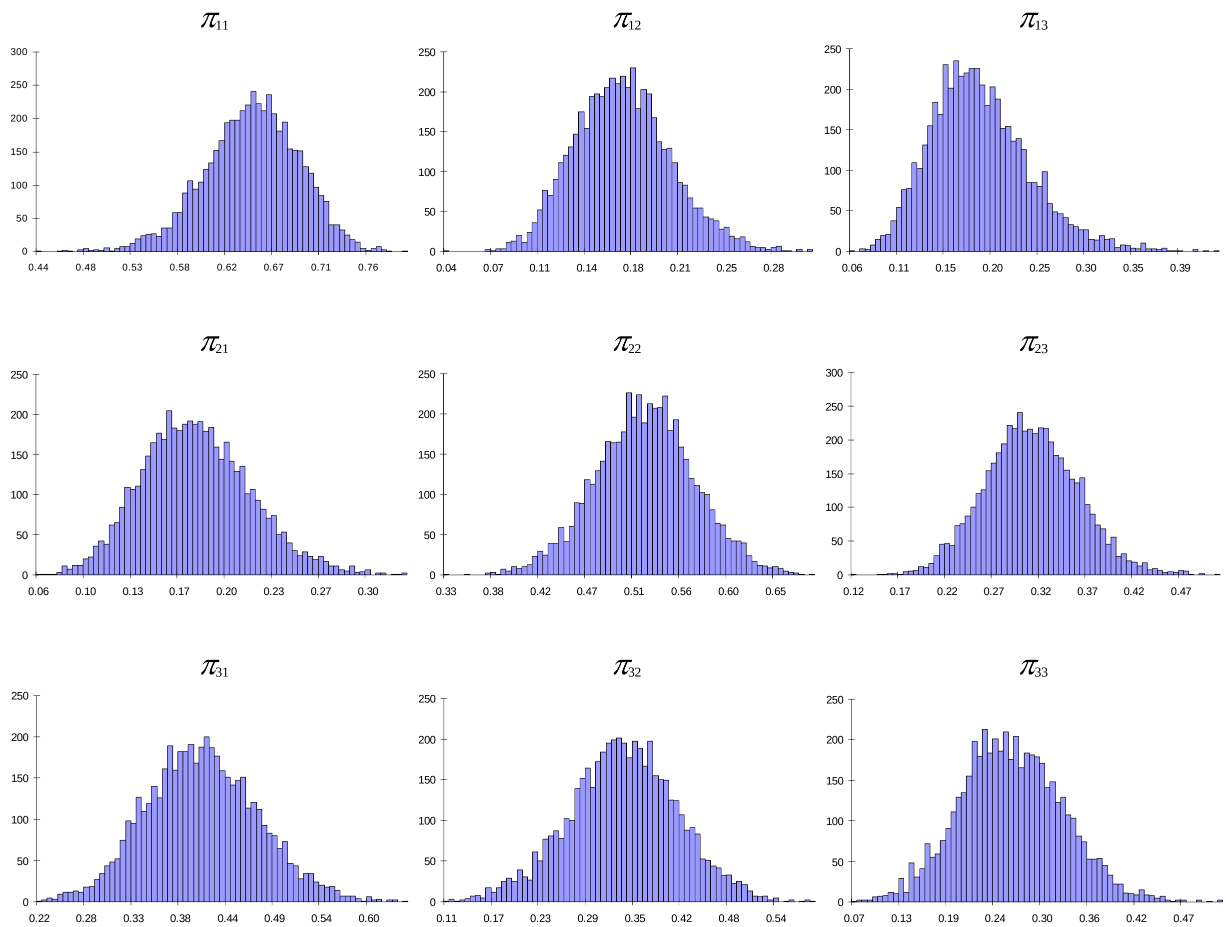
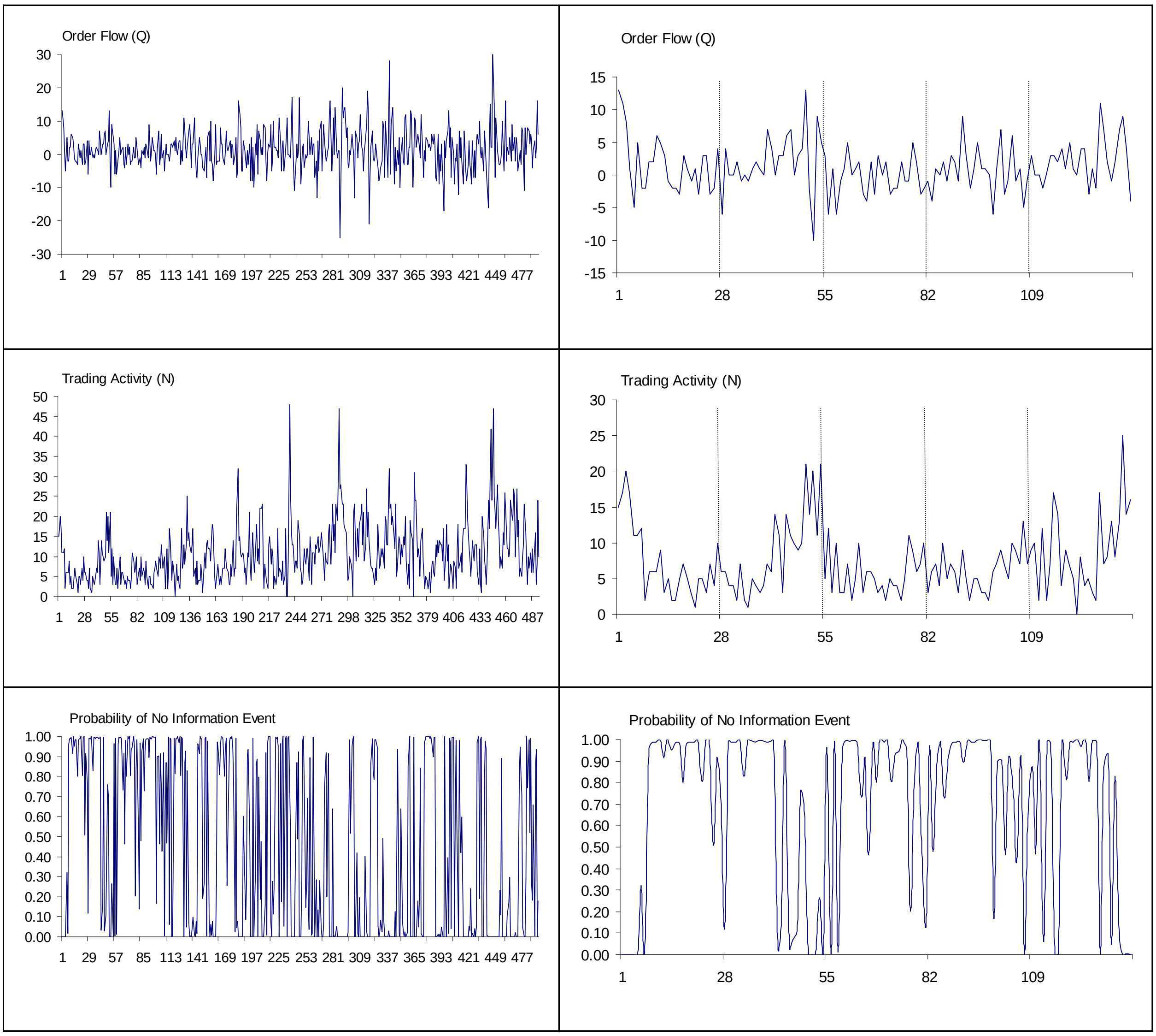


Figure 4. Net Order Flow, Trade Activity and the Posterior Probabilities for Alcoa for all trading days during the month of September, 1995.

The figures below chart the order flow variables and the corresponding posterior probabilities for each state $p(S_t = i \mid \Phi_{t-1})$. The probabilities are constructed using the Markov Mixture Model of Net Order Flow conditioning on trade direction. The left side graphs represent the set of trading intervals for the entire month of September 1995. The right side graphs are simply an enlargement of a subset of this period, which represents 1 week of trading during September. The vertical lines mark the start and end of a trading day.



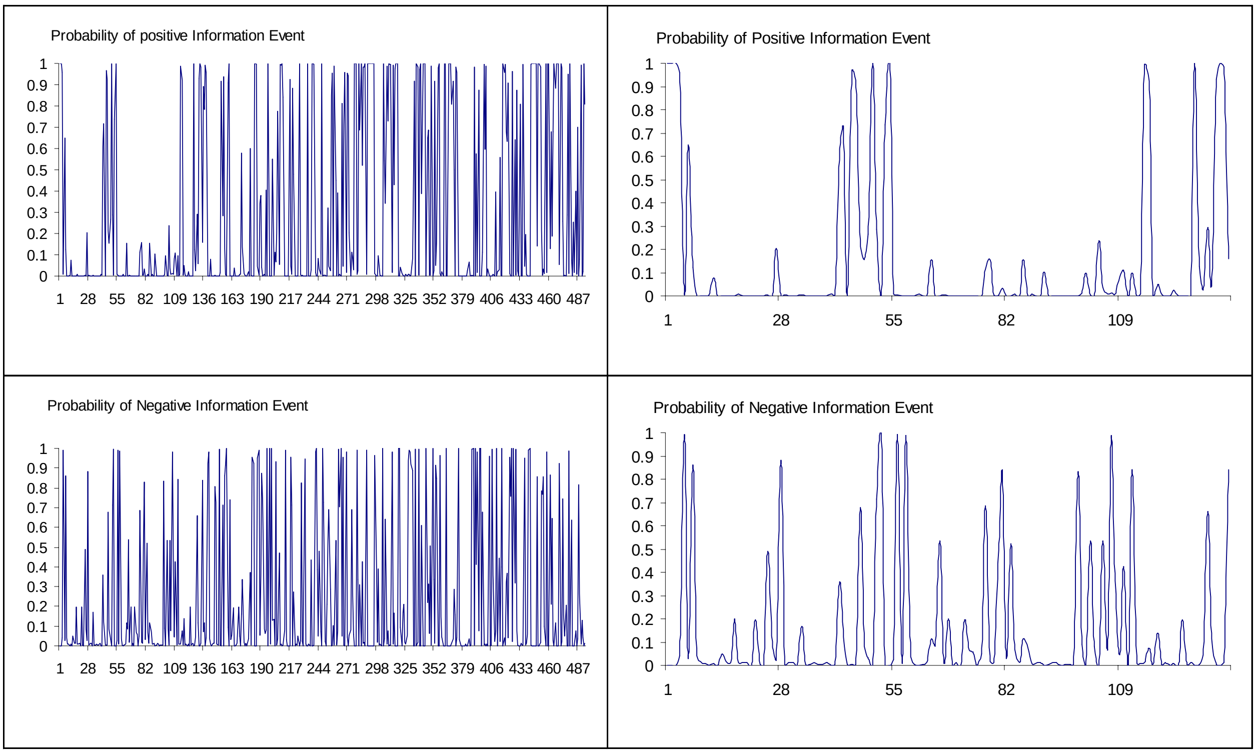

Probability of positive Information Event
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 28 55 82 109 136 163 190 217 244 271 298 325 352 379 406 433 460 487
Probability of Positive Information Event
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 28 55 82 109
Probability of Negative Information Event
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 28 55 82 109 136 163 190 217 244 271 298 325 352 379 406 433 460 487
Probability of Negative Information Event
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 28 55 82 109

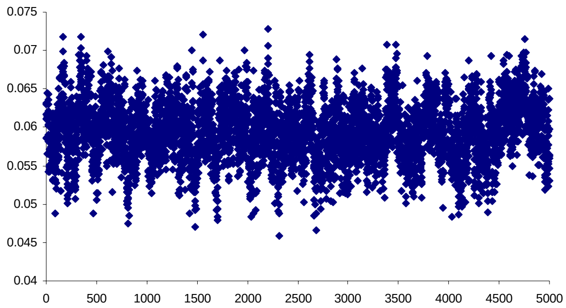
Figure 5. Posterior densities for the parameters of the Quote Setting Models using trade direction based model of Net Order Flow.

The figure is based on a Gibbs sampler run of 5000 iterations with the first 2000 discarded. Panel A reports the parameters governing the bid and ask setting functions. Panel B reports the parameters governing the expectation formation mechanism.

Panel A : Bid and Ask Equation Parameters

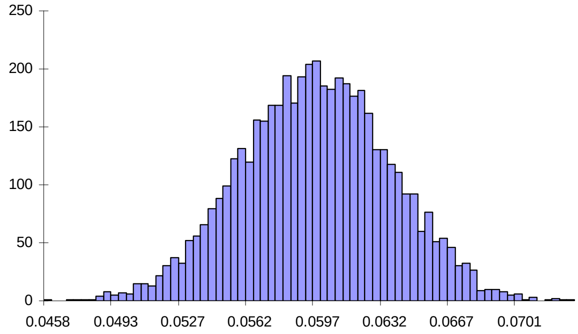
Iterates from Gibbs Run

ϕ^a

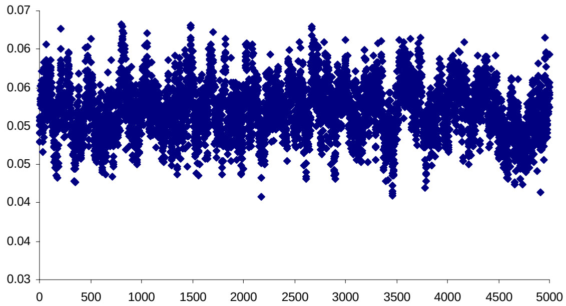


Histogram

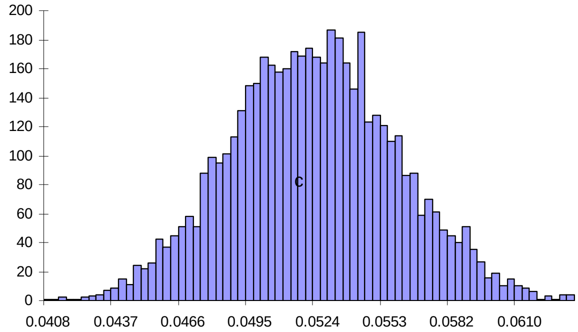
ϕ^a



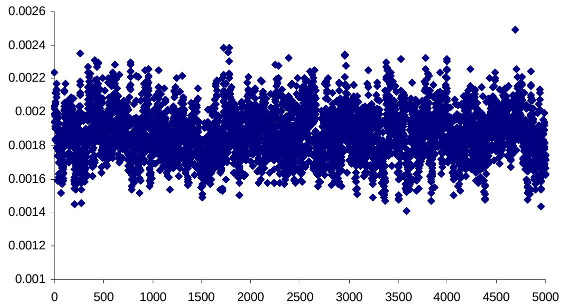
ϕ^b



ϕ^b



$\sigma^2_{u^a}$



$\sigma^2_{u^a}$

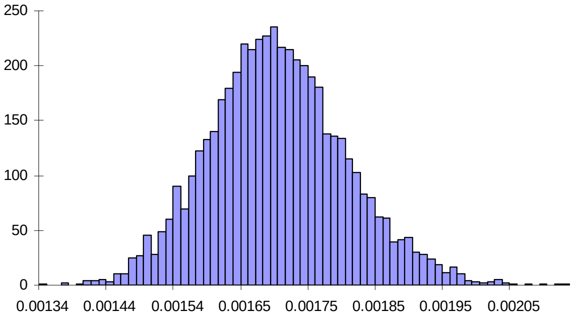
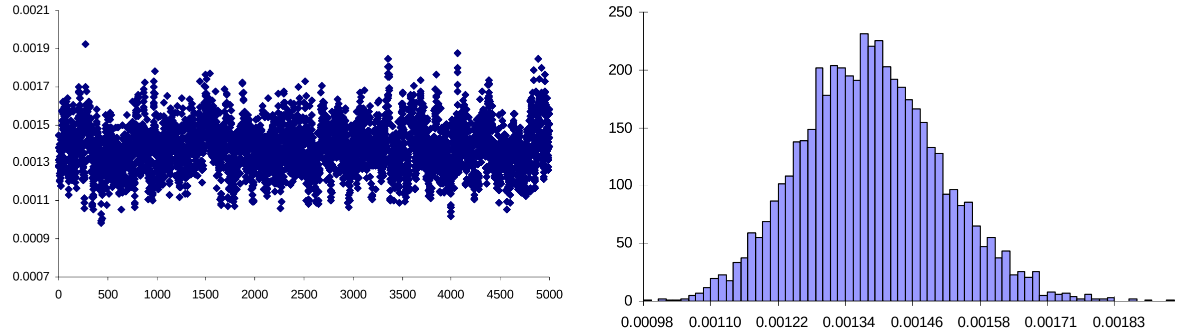


Figure 5, Panel A continued.

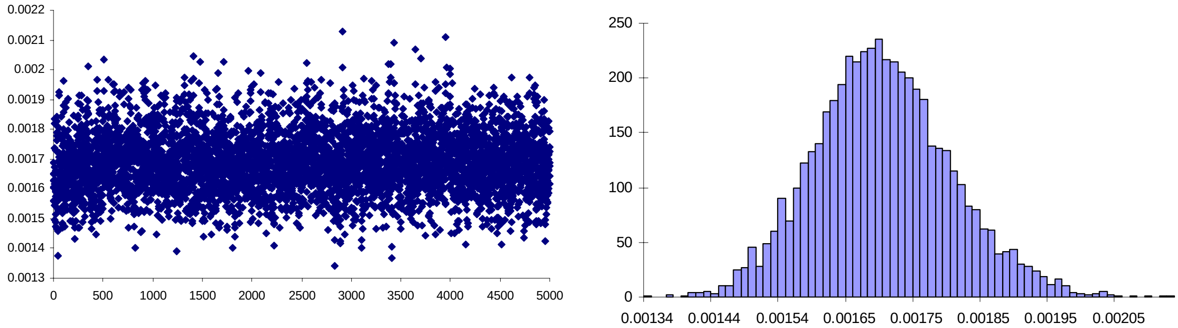
$\sigma^2_{u^b}$ $\sigma^2_{u^b}$



Panel B : Expectation Formation Parameters

Iterates from Gibbs Run *Histogram*

σ^2_{ε} σ^2_{ε}



λ_1 λ_1

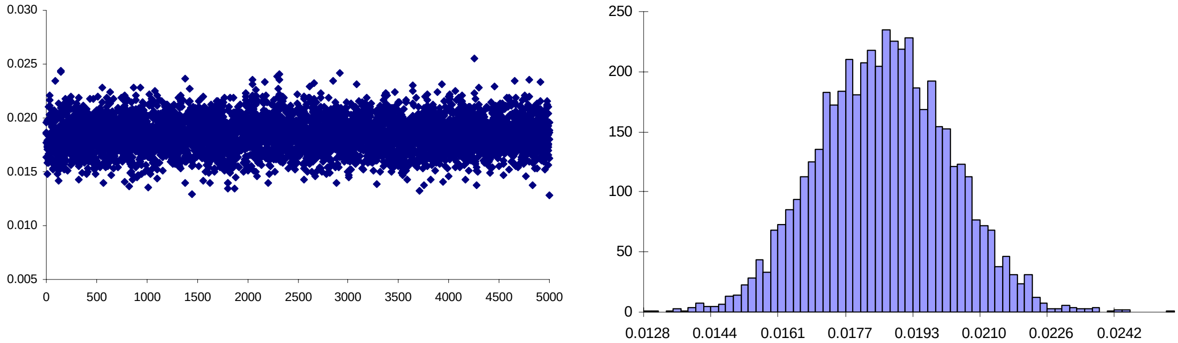


Figure 5, Panel B continued

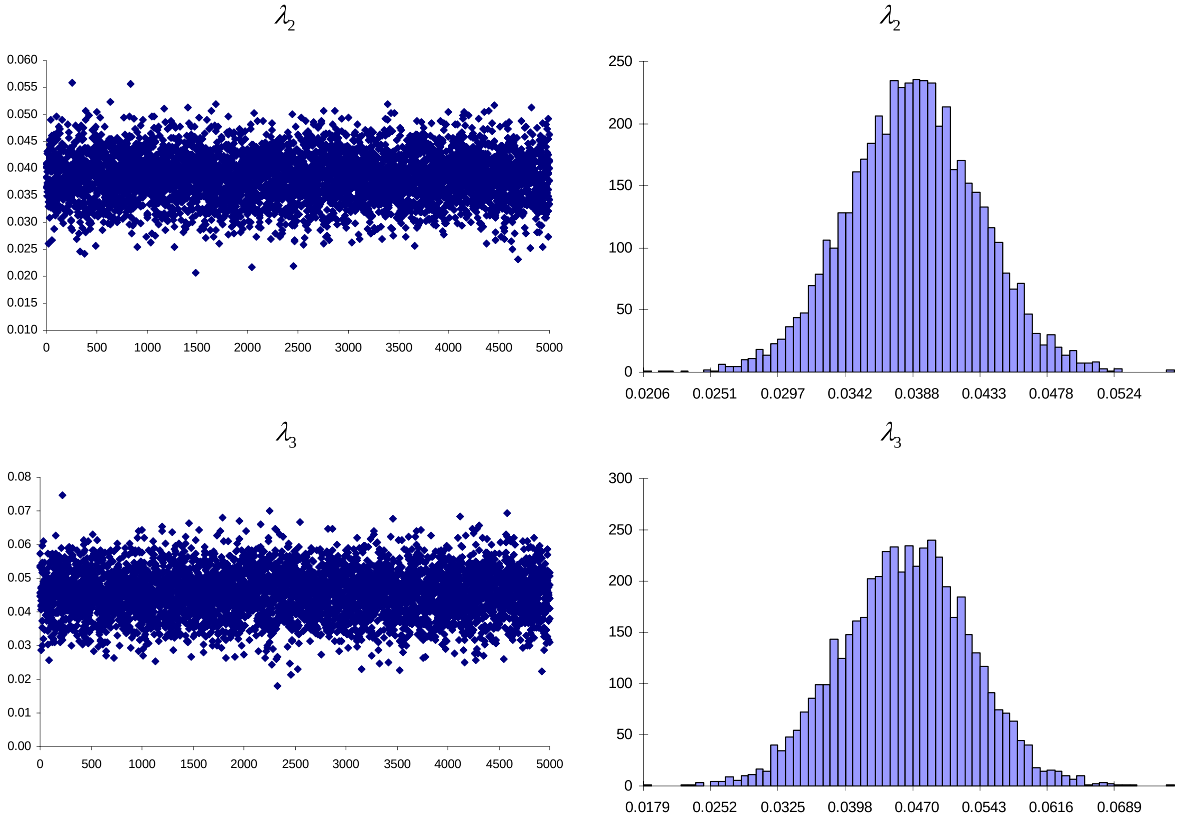


Figure 6. Efficient price series constructed from the combined Quote Setting Bounded Rational Learning Models.

The quote setting models are analysed using the parameter estimates which result from the Net Order Flow and the Bid and Ask quotes for Alcoa during 1 week of September in 1995. The first panel represents the bid and ask quotes and the bid/ask spread and the efficient price series constructed from the Kalman filter based on the parameter estimates obtained from the Gibbs sampler when using trade direction as the informative variable. The second panel charts the probability of no information over time as predicted under the trade specification. The third and fourth panels chart the actual spread and the spread attributable to information asymmetry respectively.