Liquidity in the Futures Pits:
Inferring Market Dynamics from Incomplete Data

Joel Hasbrouck

August 14, 1998

Preliminary Draft
Not for Quotation
Comments Welcome

Professor of Finance
Stern School of Business
New York University
Suite 9-190 Mail Code 0268
44 West Fourth St.
New York, NY 10012-1126

Tel: (212) 998-0310
Fax: (212) 995–4901
E-mail: jhasbrou@stern.nyu.edu
Web: http://www.stern.nyu.edu/~jhasbrou

I am grateful for financial support from a Stern Summer Research Grant.

All errors are my own responsibility.

Liquidity in the Futures Pits:

Inferring Market Dynamics from Incomplete Data

Abstract

Motivated by economic models of sequential trade, empirical analyses of market dynamics in the U.S. equities market frequently estimate liquidity from regressions of price changes on transaction volumes, where the latter are signed (positive for buyer-initiated trades; negative for seller-initiated trades). This paper estimates these specifications for transaction data from pit trading at the Chicago Mercantile Exchange. To deal with the absence of timely bid and ask quotes (generally used to sign trades in the equity market studies), this paper proposes new techniques based on Markov chain Monte Carlo estimation.

As in the corresponding equity market specifications, the model structure implies a decomposition for long-run price volatility into trade- and non-trade-related components. For the S&P contract, trades have a negligible contribution to volatility. Trades in the pork belly contract account for twenty percent of the (long-term) price volatility. Trades in the DM contract account for forty percent of the volatility. This last finding may indicate that although the futures market in the DM is dwarfed in volume by the interbank spot/forward market, the latter's relative lack of transparency causes significant price discover to occur in the futures market.

1. Introduction

The Chicago Mercantile Exchange (CME, "Merc") is a major and representative U.S. futures exchange, where most trading occurs face-to-face in physically-centralized arenas ("pits") on the Exchange floor. As a trading mechanism, the pit presently faces strong competition from electronic limit order book systems. Many of the Merc contracts, for example, are traded off-[floor]-hours on the Globex (and recently introduced Globex2) systems. In some contracts (e.g., the E-mini), floor and Globex trading occurs simultaneously. Although many observers feel that electronic trading will eventually predominate, the floor mechanism possesses a reputation for excellent liquidity and operational efficiency. Against this institutional backdrop, the present study aims at an improved characterization of the pit market mechanism.

By way of economic motivation, empirical market microstructure research seeks to measure factors thought to be important in market design and operation. Foremost among these factors are the direct costs paid by demanders of immediacy in the market and the impacts their trades have on the security price. These are often jointly (and loosely) held to summarize the market's "liquidity".

Economic models of sequential trade are, in this context, particularly important.[1] In these models a quote-setter (often a dealer or market-maker) posts bid and ask quotes; potential traders arrive one-by-one and buy or sell. After the trader has departed, the bid and ask are updated. These models are relatively mathematically tractable and flexible. Significantly (for empirical purposes), they also bear a passing resemblance to a wide range of actual markets, including the futures pits.

The empirical studies derived from these models typically estimate regressions of price changes against incoming signed order flows. A buy order, for example, is

---

[1] See Glosten and Milgrom (1985), Easley and O'Hara (1987); Easley and O'Hara (1991); Easley and O'Hara (1992a); Easley and O'Hara (1992b) and O'Hara (1995).

positively signed and hits ("lifts") the prevailing ask quote. Generally the transaction price will exceed the expected value of the security (conditional on all information available *prior* to the order arrival). This excess reflects in part a transient cost paid by the buyer for immediacy, and in part a permanent revaluation of the security. These analyses have become standard in studies of U.S. equities markets.[2] The present paper seeks to implement them for data from floor trading at the Merc.

The standard implementation of the price/signed trade regression requires both transaction (price and volume) and quote (bid and ask) data. The latter are necessary to assign a direction to the trade, usually by comparing the trade price to the quote midpoint. Bids and asks in futures pits, however, expire (unless hit) virtually instantaneously. In consequence, a contemporaneous record of these quotes comparable to that available for the equity markets does not exist.

Excellent prior studies of futures market liquidity are available. Those based on transaction-level data include Laux and Senchack (1992) and Ma, Peterson, and Sears (1992). These analyses employ return-autocovariance-based estimates of the bid-ask spread, however. This approach assumes that the direction of the trade is independent of the price movement, and cannot therefore measure informational price impacts. Manaster and Mann (1996) use Computerized Trade Reconstruction (CTR) data. These data (which are not in the public domain) establish trader identity, permit tracking of trader positions, and so support a range of interesting analyses concerning inventory control. Manaster and Mann also estimate order impacts contingent on class of trader. Identification of a buyer and seller does not, however, establish who initiated the trade (in the sense of the sequential trade models), i.e., which party hit or lifted the bid or ask exposed by the other.

---

[2] See, for example, Hasbrouck (1991a); Hasbrouck (1996a) and Madhavan, Richardson, and Roomans (1997).

It is emphasized that the absence of bid and ask data in the futures market studies is due to limitations in observation and collection procedures. Bids and asks are in fact continually being conveyed within the trading crowd. Trades occur when liquidity demanders hit these quotes. To this extent, the market structure fits the framework of the sequential trade models. But data normally essential to the estimation of these models are missing.

It is nevertheless possible to perform estimation without a complete data record, provided that one is willing to let the model structure and observed data bear the full weight of the statistical inference. In the present applications, the bid, ask and, most importantly, the direction (sign) of a given trade are viewed as latent, unobserved variables. We sign a trade, or, (more accurately) derive a probability density for the sign of the trade, conditional on the model and all observed data.

In modeling perspective, the analysis of Glosten and Harris (1988) stands as an important precursor. There, as in the present paper, transaction price and volume are observed and the order sign and efficient security price are unobserved. Glosten and Harris numerically approximate the probability density functions for these variables (conditional on the observed data) and the likelihood function for the observations. Estimation proceeds by maximum likelihood. The analysis falls within the general approach to nonlinear state-space model described by Kitagawa (1987). Glosten and Harris apply the technique to a sample of NYSE transaction data.

The empirical model in the present paper generalizes on Glosten-Harris in allowing a more flexible treatment of discreteness, clustering and trade-price impacts. A more fundamental difference, however, lies in estimation methodology. The present paper employs a Markov chain Monte Carlo (MCMC) estimator, the Gibbs sampler, which is attractive both analytically and computationally.[3]

The models and methods presented here also differ from the usual empirical market microstructure analyses in that they are cast in a Bayesian framework. Bayesian methods are usually employed to incorporate prior beliefs about model structure or parameters. Indeed, this consideration has not received proper attention in market microstructure. For example, despite our strong priors that the bid-ask spread is positive, moment-based sample estimates using the Roll (1984) procedure are frequently "negative" (or undefined). But the more compelling motivation for the use of Bayesian methods here lies in the analytical and computational ease with which latent variables (such as the unobserved trade direction) may be incorporated.

The paper begins with a summary of the trading procedure and descriptive statistics of the price and volume series. This serves to establish features of the data that arise in modeling. The paper then turns to issues of modeling, estimation and economic interpretation. This part of the paper is organized around a series of models of increasing richness and complexity, beginning with a reworking of the Roll model (Section 3), and continuing through models that incorporate discreteness (Section 4), clustering (Section 5) and asymmetric information (Section 6). The sequential presentation of the models serves to illustrate Bayesian modeling and estimation principles that are, although well-established and standard in other contexts, relatively unfamiliar in empirical microstructure. The reader uninterested in methodology may prefer to skip these sections (3-6). A comprehensive model is presented and estimated in Section 7. A brief summary concludes the paper in Section 8.

---

[3] Useful introductory references in this field include Gilks, Richardson, and Spiegelhalter (1996) (for a concise overview of MCMC techniques), Casella and George (1992) (for the Gibbs sampler) and Chib and Greenberg (1996) (for applications in econometrics).

2. Data overview

This section describes the futures transaction data with a view to illuminating the distinctive features of the data that an empirical model should account for, or at least accommodate.

The Chicago Mercantile Exchange is a major U.S. futures exchange. Their web site (at www.cme.com) provides a comprehensive description of the Exchange, instruments, trading mechanisms and data (including that used in the present study). The trading arrangements at the CME are typical of U.S. futures exchanges. Traders interact face-to-face on the exchange floor. They compete by shouting and signing acceptable price/trade combinations. Thus, bids and offers are transient, options that vanish unless exercised immediately. They are frequently refreshed, as a trader may continually repeat a bid or offer. But unlike the U.S. equity markets, there is no presumption that a bid or offer is good until explicitly canceled or modified. This transience does not, however, invalidate the sequential trade framework, since we are still in a world where the quote setter moves first and the (potential) “market order” trader follows.

An observer on the floor sees bids, offers and trades. In real time, however, off-floor participants must rely on the electronically disseminated tick data. The reported price is the most current trade price. This is updated only when a trade at a new price occurs. This differs, of course, from the last sale reporting practices in U.S. equities markets, wherein a trade is reported even if it is at the same price as the previous trade. Smith and Whaley (1994) discuss estimators of the Roll bid-ask spread using time and sales data.

The data used in the present study, however, are drawn from the CME’s volume-tick files. These data consist of time-stamped trade prices and volumes, i.e., a record essentially similar to what one receives from the U.S. equity market’s Consolidated Transaction System. These data are synthesized (based in part on audit trail data) after the close of trading, however, and are not available to market participants in real time. The sample is drawn from the volume-tick files for the first full two weeks of January 1998.[4] This section summarizes characteristics of sixteen heavily-traded contracts. In subsequent sections, detailed time series analyses are described for three representative contracts (pork bellies, the Deutschemark and the S&P composite index).

Table 1 describes various features of the analyzed contracts. Of particular relevance for the paper are the tick sizes. As a proportion of the contract price, they are often dramatically lower than those commonly encountered in equity markets. A tick of $1/16 is 0.125% of a $50 stock. This is at least an order of magnitude greater than that of any of the futures contracts.

Table 2 reports trade characteristics. These suggest the scale and timing of the transactions. For sheer pace of trading activity, the S&P composite contract stands out. It exhibits an *average* intertransaction time of only four seconds. Trades frequently occurred within the same second. The economic framework of the sequential trade models generally assumes that trade reports are instantaneously disseminated and evaluated. In the S&P index pit, at least, an individual trader's information set is unlikely to be this current.

Table 3 reports standard deviations and first-order autocorrelations of the intertransaction returns. The returns are measured alternatively as difference in log price and as difference in price level (measured in ticks). The results for the level prices suggest that intertransaction return volatility is not large relative to the tick size, and therefore that discreteness may be an important consideration in assessing a contract's volatility. The first-order autocorrelations are negative, presumably reflecting bid-ask bounce. This feature is discussed more completely in Section 3.

---

[4] The data files used are those from the CME's website with a prefix "vt" (for volume-tick).

A phenomenon closely related to discreteness is clustering, the tendency of trades (and presumably quotes) to cluster on "natural" multiples of the minimum tick. There are various ways of describing clustering. In the NYSE data examined by Harris (1991), the $1/8 tick clearly motivates "two-based" clustering on whole numbers, halves and quarters. The tick size across futures contracts, however, is not uniform. A preliminary look at the data suggests that more generality is needed. Accordingly, this study examines both two- and five-based clustering, i.e., the incidence of prices that fall on κ-multiples of the minimum tick, where $\kappa = 2^i 5^j$ for small nonnegative integers i and j. Letting $0 \le i \le 3$ and $0 \le j \le 2$ generates a set of possible values for κ: $\kappa \in \mathrm{K} = \{1, 2, 4, 5, 8, 10, 20, 25, 40, 50, 100, 200\}$.

We invoke no economic or mathematical laws here. This approach merely seems to give rise to a set of numbers that many people would regard as natural or convenient. The method of construction does imply, however, an (incomplete) ordering of its members. For example, if traders are observing an implicit tick size of κ=5, and seek to establish a coarser tick, the "natural" choice is κ=10 (rather than κ=8). In this sense, all κ>1 have a finer predecessor.[5]

Two descriptive measures are useful here. Let f_κ be the sample frequency of prices that lie on a κ-multiple. The clustering frequency is simply the excess above expectation: $f_\kappa^C = f_\kappa - (1/\kappa)$. It is also helpful to describe the incremental change in clustering associated with moving from a finer to coarser κ (in the sense of the ordering above). For example, clustering on multiples of two will elevate the incidence of four-multiples. The incremental clustering is defined in such a way as to correct for clustering at the next finer level. That is, $\Delta f_4^C = f_4 - f_2/2$, and so on.

---

[5] More precisely (letting "$\succ$" denote "is preceded by"): $200 \succ 100$, $100 \succ 50$, $50 \succ 25$, $40 \succ 20$, $25 \succ 5$, $20 \succ 10$, $10 \succ 5$, $8 \succ 4$, $5 \succ 1$, $4 \succ 2$, $2 \succ 1$. As a formalism, the predecessor for κ is Max[κ/2,κ/5] restricted to the set K; the ordering is incomplete because we can't assert $8 \succ 5$ or $5 \succ 8$.

Table 4 reports clustering frequency percentages for the sample contracts; incremental clustering frequencies are given in Table 5. The interaction between the two measures may be illustrated for the feeder cattle contract. The raw frequencies (not reported in the table) are f_2=76% and f_4=39%. In a large random sample, we would expect 50% of prices to lie on even tick-multiples, so the clustering percentage at κ=2 is $f_2^C = 26\%$. Given that 76% of prices lie on two-tick multiples, we would expect 38% to lie on four-tick multiples. The observed frequency exceeds this by 1% (the incremental clustering Δf_4^C for this contract).

Table 4 implies that many (but not all) of the contracts exhibit clustering. Clustering at multiple levels is in most cases a consequence of clustering at a relatively fine level. (Compare, for example, the clustering and marginal clustering percentages for pork bellies and feeder cattle.) A notable exception is the Nasdaq 100 contract, for which there is pronounced marginal clustering at ten-multiples.

Economic explanations for clustering are varied. Harris (1991); Harris (1994) suggests that negotiating parties may adopt a supra-minimum tick convention as a device for reducing the number of rounds of bargaining, and therefore the bargaining cost. In this view, the coarser rounding randomly works for and against a trader, tending to average out to zero. Traders economize on their bargaining time. This cost savings is presumably passed on to off-floor traders via competition in the market for broker services.

It is also suggested, however, that when there are barriers to entry in the provision of liquidity services, clustering may serve as an implicit collusive coordination mechanism (see Kandel and Marx (1997) and Dutta and Madhavan (1997)). This has been most strongly alleged for the Nasdaq dealer market prior to the recent reforms.[6]

---

[6] The literature on clustering at Nasdaq is large. Key references include Christie, Harris, and Schultz (1994); Schwert (1997).

Although the trading mechanism in the futures markets differs profoundly from the dealer market structure of Nasdaq, there are also some striking similarities. In their study of the CTR data, Manaster and Mann (1996) note that "The most frequent combination is a customer order . . . filled by a market maker . . . or local." (p. 956). Essentially, most liquidity seems supplied by dealers or quasi-dealers, rather than outside customers. Furthermore, the Nasdaq order preferencing arrangements that keep order execution within a subset of dealers are mirrored by the futures exchanges' Brokers' Associations (U. S. Commodities Futures Trading Commission (1997)). This paper does not attempt to resolve the underlying reasons for clustering, however, which remain an important topic for further research.

As a guide for modeling strategy, the analysis suggests to this point the following considerations. First, motivated by the economic sequential trade models, it seems desirable (as in the equity market studies) to allow for trade-driven price impacts of both a transient (cost-related) and permanent (informational) nature. The results of this section suggest that in addition, discreteness is important because the tick size is generally on the same scale as intertransaction volatility. Transaction prices furthermore exhibit a tendency to clustering that is for some contracts highly pronounced. With these features in mind, we proceed to discuss a series of empirical specifications.

3. The Roll model of the spread

Roll (1984) presents a model that is, by reason of its simplicity and ease of implementation, a useful starting point. In this model, transaction prices behave as a random-walk-plus-noise, wherein the random-walk is the efficient price of security and the noise is "bid-ask bounce". Throughout this paper, the term "efficient price" will be used in a sense common to the sequential trade models, the expected terminal value of the security conditional on all public information (including the trade history).

a. The basic model

A variant of the Roll model is as follows. Let the efficient price be denoted M_t. Its logarithm $m_t = \log(M_t)$ is assumed to evolve as a random walk:

$$m_t = m_{t-1} + u_t \tag{1}$$

where the u_t are zero-mean innovations stemming from the arrival of new public information. The (log) bid and ask prices are given as

$$\begin{aligned} b_t &= m_t - c \\ a_t &= m_t + c \end{aligned} \tag{2}$$

where c is the half-spread, presumed to reflect the quote-setter's cost of market-making.

The direction of the incoming order is given by the Bernoulli random variable $q_t \in \{-1,+1\}$, where -1 indicates an order to sell (to the quote-setter) and +1 indicates an order to buy (from the quote-setter). That is, q_t specifies which side of the trade is, in terminology sometimes encountered, the trade initiator or trade aggressor. In security markets we usually assume the Bernoulli outcomes equally likely, which implies a symmetry in the order flow. For expositional convenience, it will be assumed throughout that the unconditional probabilities of either q_t outcome is one-half. This restriction can easily be relaxed, and there are some markets (e.g., the real estate market) where it would not be appropriate.

Most implementations of the Roll model assume that q_t is independent of $\Delta m_t = u_t$, i.e., that the direction of the trade is independent of the efficient price movement. This rules out the asymmetric information aspects of the sequential trade models, and so is not an innocuous assumption. For the sake of expositional clarity, we initially adopt the assumption. But it is not, in this paper's approach, an essential requirement, and is relaxed in Section 6. Depending on q_t, the (log) transaction price is either at the bid or the ask:

$$p_t = \begin{cases} b_t \text{ if } q_t = -1 \\ a_t \text{ if } q_t = +1 \end{cases} \tag{3}$$

We typically possess a record of trade prices and seek to estimate the two model parameters.

Estimation usually proceeds via method of moments. The model implies a variance and first-order autocovariance for the log price changes of:

$$\begin{aligned} Var(\Delta p_t) &= \sigma_u^2 + 2c^2 \\ Cov(\Delta p_t, \Delta p_{t-1}) &= -c^2 \end{aligned} \tag{4}$$

The corresponding sample moments imply estimates for σ_u^2 and c that possess all the usual properties of GMM estimators, including consistency and asymptotic normality.

The model forces the first-order autocovariance to be nonpositive irrespective of the sign of c. In sample data, however, this property is often violated. In his examination of U.S. stock data, for example, Roll finds that autocovariance estimates based on 21 daily returns are positive roughly half the time. In discussing the sampling properties of this estimator, Harris (1990b) concludes that noise in typical applications will frequently lead to positive autocovariances, even if the model is correctly specified.

b. Bayesian estimation

Our conviction that the spread must be positive is a prior belief, and as such is most naturally incorporated in a Bayesian framework. To illustrate, we augment the model with a distributional assumption: $u_t \sim N(0, \sigma_u^2)$. The model parameter set is $\Theta = \{\sigma_u^2, c\}$. Denote the prior parameter density as $\pi(\Theta)$. We seek the parameter posterior $f(\Theta|p) = f(p|\Theta)\pi(\Theta)/f(p)$ where $p = \{p_1, \ldots, p_T\}$ is the vector of observed prices.

Direct evaluation of this posterior is beset by difficulties from the outset. The data likelihood function $f(p|\Theta)$ involves the unobserved $q = \{q_1, \ldots, q_T\}$. No tractable closed-form representation exists, and one is left with numerical approximations (Harris (1990a)). Incorporating the parameter prior complicates matters still further.

Surprisingly, the situation is simplified if we bring the unobserved (latent) variables into the problem explicitly, writing the posterior as a joint conditional distribution over these latent variables as well as the parameters. If we possessed this posterior, $f(\Theta, q|p)$, we could obtain the parameter posterior by integrating out *q*. It is not immediately clear, however, that this would help matters. The expanded posterior $f(\Theta, q|p)$ has substantially greater dimensionality and complexity than the one we originally sought $f(p|\Theta)$. Nor does the task of integrating out *q* (to obtain the parameter posterior) appear trivial.

The MCMC approach neatly solves both problems. In the first place, it works with simulated samples. If we possess a sample of draws from $f(\Theta, q|p)$ denoted $\Theta^{(i)}, q^{(i)}$ for $i = 1, \ldots, N$, then we can view the $\Theta^{(i)}$ as originating from the marginal $f(\Theta|p)$. (Essentially we integrate out the *q* by discarding them.) These draws can be used to characterize all features of the posterior distribution.

Sampling, in turn, is facilitated by the Gibbs principle, wherein a draw from a complicated joint pdf is built by cycling over (simpler) conditional pdfs. The steps in this sampling are:

0. Initialize *q*.
1. The conditional parameter draw: draw Θ from $f(\Theta|p, q)$.
2. The conditional latent variable draw: draw *q* from $f(q|p, \Theta)$.

We then generate the required sequence of draws by iterating between steps 1 and 2. We discuss the details of these conditional draws below, but it is first useful to summarize a few general considerations.

Firstly, the technique of treating latent data and parameters equivalently is a hallmark of modern Bayesian analysis (Tanner (1996)). In this problem the latent data arise naturally in the structural model, but in some situations latent variables of a more artificial nature may be introduced as simplification devices. As in the present instance, however, this approach greatly expands the dimensionality of the problem, necessitating numerical methods substantially more powerful than the numerical integration computations traditionally used. Markov chain Monte Carlo (MCMC) methods (including the Gibbs sampler) have proven particularly useful in this respect.

The steps enumerated above describe a Gibbs sampler used at the top level of the problem, iterating between parameters and latent data. As will shortly be seen, however, Gibbs samplers are used more pervasively here, *within* the parameter and latent data simulations. In fact, whenever possible, I rely in this paper on single-step Gibbs samplers, in which each conditional draw is of a single random variable (parameter or latent datum). Among the various possibilities, these are usually the most tractable analytically and most amenable to exposition (although they are frequently not the most computationally efficient). Finally, although the Gibbs sampler is valid under fairly general conditions, these must not be taken for granted. It will be seen that even for some of the simple models considered in the present paper, certain single-step Gibbs samplers that initially appear attractive fail. Fortunately in these situations, there are usually tractable alternatives (most commonly involving joint conditional draws).

This model (and all models discussed in this paper) can be written as nonlinear (and non-Gaussian) state-space models. Carlin, Polson, and Stoffer (1992) proposed estimating such models with Gibbs samplers that single-step through time. All of the estimation strategies discussed in this paper are of this form. (The joint draws used in some models are joint across state variables at a single point in time.) The nonlinear, non-Gaussian framework is sufficiently general to encompass many modifications for which single-step samplers may be computationally inefficient. In such situations, when a model can be expressed in a fashion that retains a high degree of Gaussian structure, the blocked samplers proposed by Shephard (1994) and Carter and Kohn (1994) may be more useful. Manrique (1997); Shephard and Manrique (1997); Manrique and Shephard (1998) discuss estimation when the observations are discrete.[7]

Conditional parameter draws.

The model implies

$$\Delta p_t = c\Delta q_t + u_t \, , \tag{5}$$

i.e., a regression in which *c* appears as a coefficient. This is easily handled within a Bayesian linear regression framework. (See the summary in the appendix.) It is convenient to apply the Gibbs principle again, breaking step 1 above into

1a. Draw σ_u^2 from $f\left(\sigma_u^2 \middle| c, p, q\right)$

1b. Draw *c* from $f\left(c \middle| \sigma_u^2, p, q\right)$

Note that it is not necessary to cycle between 1a and 1b "to convergence" before moving on to step 2. It is convenient to use a prior $\pi(c) = N^+\left(\mu, \sigma^2\right)$, where $N^+\left(\mu, \sigma^2\right)$ the normal density with mean μ and variance σ^2 truncated to the nonnegative real line. (Note that μ and σ^2 serve as formal parameters only: the mean and variance of the truncated density are not μ and σ^2.) A convenient prior for the variance parameter is the inverse gamma.

Conditional latent data draws.

To execute the conditional latent draws, we invoke the Gibbs principle again, making successive draws from the distribution defined by $\Pr\left[q_t \middle| q_{/t}, p, \Theta\right]$ where $q_{/t} = \{q_t, \ldots, q_{t-1}, q_{t+1}, \ldots q_T\}$. All of the Gibbs samplers described in this paper proceed in this fashion.

---

[7] The evolution of approaches to MCMC estimation of stochastic volatility models is a useful guide here. See Shephard (1993); Jacquier, Polson, and Rossi (1994); Shephard (1994); Shephard and Pitt (1997b); Kim, Shephard, and Chib (1998).

Note first that the structure of the model implies that $\Pr\left[q_t \middle| q_{/t}, p\right] = \Pr\left[q_t \middle| m_{t-1}, m_{t+1}, p_t\right]$. (When discussing the data draws, we will notationally suppress the parameter vector as a conditioning argument.) From Bayes rule, $\Pr\left[q_t \middle| m_{t-1}, m_{t+1}, p_t\right] \propto \Pr\left[p_t \middle| m_{t-1}, m_{t+1}, q_t\right] \Pr\left[q_t \middle| m_{t-1}, m_{t+1}\right] = \Pr\left[p_t \middle| m_{t-1}, m_{t+1}, q_t\right](1/2)$. Furthermore, $\Pr\left[p_t \middle| m_{t-1}, m_{t+1}, q_t\right] \propto f\left(m_t \middle| m_{t-1}, m_{t+1}\right)$ where the latter is evaluated at $m_t = p_t - cq_t$. The conditional distribution of m_t is:

$$\begin{aligned} f\left(m_1 \mid m_2\right) &= N\left(m_2, \sigma_u^2\right) \\ f\left(m_t \middle| m_{t-1}, m_{t+1}\right) &= N\left(\frac{m_{t-1} + m_{t+1}}{2}, \frac{\sigma_u^2}{2}\right) \quad \text{for } t = 2, \ldots, T-1 \\ f\left(m_T \middle| m_{T-1}\right) &= N\left(m_{T-1}, \sigma_u^2\right) \end{aligned} \tag{6}$$

To proceed, we compute $f\left(m_t \middle| m_{t-1}, m_{t+1}\right)$ at $m_t = p_t - cq_t$ for $q_t \in \{-1, +1\}$, normalize to obtain the l.h.s. probabilities, and make the Bernoulli draw of q_t.

Identification

In discussing the conditional coefficient draws, we imposed a prior belief of nonnegativity on the half-spread parameter c. While economically reasonable and statistically desirable, this restriction (or something like it) is actually required to identify the model. This requirement derives from the two related conventions involving the signs of q_t and c. It is customary to take $q_t = +1$ and –1 to indicate incoming buy and sell orders and $c>0$ to reflect the cost of market-making. Yet since the two quantities appear only in the product cq_t, the model is observationally equivalent to one in which the order signing is reversed (e.g., purchases are signed negatively) and $c<0$. Without the restriction that $c>0$, the posterior draws will (in time) cycle over both possibilities. We would expect to see posteriors (both for c and the q_t) symmetric about zero (and quite possibly bimodal). Simulations verified that this indeed occurs. The analyst examining only the posterior means or medians would conclude that the data possessed no power in assigning trade direction and that trading costs were zero.

Monitoring and summarizing the estimation process

The output from the Gibbs sampler described above is a sequence of parameter estimates $\Theta^{(i)}$, for $i = 1, \ldots, N$. where *N* is the number of simulations. After the process has mixed sufficiently that the influence of the starting values is negligible (the "burn in" period), this sequence may be viewed as a set of dependent draws from the desired posterior. The matters of judgement here involve deciding the size of *N* and the length of the burn-in period (i.e., portion of the initial *N* draws to be discarded). There are no universally accepted tests here, but all sources recommend visual (graphical) assessment of the individual parameter draws. In summarizing the parameter posteriors, we report here means, together with standard errors corrected for serial dependence, and standard deviations.

It is often convenient to discuss transformations of the model parameters (σ_u rather than σ_u^2, for example. It is easy to analyze the posterior distribution of an arbitrary transformation, say $\theta = F(\Theta)$, since the sequence $\theta^{(i)} = F(\Theta^{(i)})$ is a sequence of draws from this distribution.

Model Comparisons and Specification Analysis

Model comparison in a Bayesian framework is performed using the posterior odds ratio, or (as the number of observations becomes asymptotically large) the Schwarz information criterion. The necessary numerical likelihoods could be computed using the auxiliary particle filter approach of Shephard and Pitt (1997a). These are used in Hasbrouck (1998b), but have not yet been implemented for the present models.

Although the models discussed in this paper do not in general possess "estimated residuals", it may be useful to examine related quantities. In forming parameter estimates with the Gibbs sampler described above, we "discarded" the simulated latent data, the $q_t^{(i)}$ for t=1,…,T and i=1,…,N. In fact, these are in principle drawn from the smoothed distribution: $q_t^{(i)} \sim f\left(q_t \middle| p_1, \ldots, p_T\right)$, and the average (over *i*) $\hat{q}_t = \sum q_t^{(i)} \big/ N$ is an estimate of the expectation of this smoothed distribution.

Smoothed quantities are useful in assessing the ability of the estimation procedure to recover latent data. In the present model, the trade direction variable has the property that $\mathrm{Var}(q_t) = 1$. If the data and model parameters are such that the smoothed estimates $\hat{q}_t$ are always close to –1 or +1, we would conclude that that the model and data are signing the trades reliably. With a little more structure, we can derive a more intuitive measure. Let π^C be the probability that a trade is correctly classified on a given draw, i.e., $\pi^C = \Pr\left[q_t = +1 \middle| q_t^{(i)} = +1\right] = \Pr\left[q_t = -1 \middle| q_t^{(i)} = -1\right]$. Given the assumed symmetry of the trades, $\pi^C = \Pr\left[q_t^{(i)} = +1 \middle| q_t = +1\right] = \Pr\left[q_t^{(i)} = -1 \middle| q_t = -1\right]$. It follows that

$$\begin{aligned} E\hat{q}_t^2 &= \frac{1}{2}\left(E\left[\hat{q}_t^2 \middle| q_t = +1\right] + E\left[\hat{q}_t^2 \middle| q_t = -1\right]\right) \\ &= \frac{1}{2}\left(\left[\pi^C(+1) + \left(1-\pi^C\right)(-1)\right]^2 + \left[\pi^C(-1) + \left(1-\pi^C\right)(+1)\right]^2\right) \\ &= \left(2\pi^C - 1\right)^2 \end{aligned} \tag{7}$$

This in turn suggests that π^C may be estimated by

$$\hat{\pi}^C = \left(1 + \sqrt{Var(\hat{q}_t)}\right)\big/2 \tag{8}$$

It must be emphasized that this estimate measures the classification accuracy only in a conditional sense, assuming that the model is correctly specified. It is not an overall measure of model fit. At best it suggests the ability of the model and data to support inference about a key latent variable.

c. Application to the futures market data

The model discussed in this section is naïve in many respects, some of which will be remedied later. Its tractability nevertheless recommends it as a starting point for analysis. First, by way of preliminaries, Table 6 reports moment estimates of the Roll model for each contract. In all cases, the sample first-order return and price change autocovariances are negative, so the problem of imputing a negative or undefined spread mentioned above does not arise.

The model discussed above was estimated for the three representative contracts for the first 1,000 trades in the sample using 10,000 draws of the Gibbs sampler.[8] Convergence and mixing were monitored visually. While space considerations preclude full presentation of these results, the graphical output for the pork belly contract (Figure 1) is typical. For both parameters, there are three graphs. The left-most graphs depict the draws themselves, which evince no obvious large persistent components. The autocorrelations of the draws are low (middle graphs) and the histograms of the draws (right-most graphs) suggest well-defined densities.

Table 7 presents summary statistics for the (simulated) model posterior distributions. We focus here primarily on the posterior means of the *c* (log-half-spread) parameters. In comparing these estimates with the moment-based estimates in Table 7, one might expect substantial agreement, since both are based on the same model and sample. Differences exist, however (most strikingly for the pork belly contract).

The discrepancies in the estimates appear to arise from differences in how the two approaches use sample information. For example, with the moment approach, all of the negative autocovariance is attributed to the trading cost parameter *c*. In the Gibbs approach, *c* is a coefficient in a regression that assumes (but does not enforce) independence of residuals. To further explore the sources of differences would serve little purpose, however. Construction of a more realistic more stands as a more pressing concern.

[8] The models in this paper were investigated with both diffuse proper priors and noninformative (frequently improper) priors. Although it is generally impossible to verify that improper priors lead to proper posteriors in Gibbs samplers, the results in the present case were substantially similar. For brevity, only the results based on the noninformative priors are reported.

d. Further perspectives on signing trades

Given the importance attached by the sequential trade models to order direction, it is not surprising that this has arisen as a perennial concern in microstructure modeling. In the NYSE's unusually-detailed TORQ dataset (Hasbrouck (1992); Hasbrouck (1996b)), it is possible to associate many trades with the actual underlying orders. More commonly, however, trade direction is inferred from related price data. As noted in the introduction, the usual practice is to sign trades by reference to the prevailing quotes (see Hasbrouck and Ho (1987), Hasbrouck (1988), Lee and Ready (1991) and Odders-White (1997)).

Consider the following hypothetical analysis. In a sample of transaction price data, we adopt the following rule. If a price p_t occurs on an uptick (or zero-uptick), we set q_t=+1; on a downtick (or zero-downtick), we set q_t=–1. We then estimate equation (5) using the constructed q_t as regressors. This analysis is highly improper because our classification rule induces correlation between measurement errors in q_t and the model disturbance. Yet in the present framework, we seem to be drawing inferences about trade direction that are very similar. A pattern of successive price upticks, for example, will be viewed as a procession of "buy" orders. It might therefore appear that the present analysis falls to the same objections as the proposed naïve one.

There are, however, two crucial differences. First, the present procedure does not assign to a trade a single direction that is used in all subsequent computations. Instead, it imputes a probability density over both (buy and sell) alternatives. In this sense, the procedure explicitly models the measurement error (uncertainty) concerning trade direction. In the second place, the trade directions and model parameters are estimated jointly. This essentially allows uncertainty about model parameters to affect uncertainty about trade direction. We are still, of course, assuming that the model is correctly specified. But we do not assume "full knowledge" (i.e., correct parameter estimates) of the model in the process of assigning trade direction.

e. Random costs of market making

The model discussed to this point assumes a constant half-spread c. Because market conditions are likely to be changing, it is a useful generalization to permit this cost to vary randomly. This modification corresponds to letting c in equation (2) be replaced by c_t, an iid nonnegative random variable. This modification adds T new latent random variables: $c_t, t = 1,\ldots,T$. One might be tempted to modify the above Gibbs strategy by reasoning as follows. At the conditional latent data draw step we need to simulate the c_t, but this simulation consists of a trivial calculation: given p_t, m_t and q_t, $c_t = (p_t - m_t)/q_t$ (an identity). We then proceed to draw q_t as previously described. A moments reflection will confirm, however, that with this procedure c_t can't move.

There are two ways to remedy matters. First, we can replace the improper sequential Gibbs draw with a (joint) draw from the joint conditional distribution $f(m_t, c_t, q_t | p_t, \ldots)$. This is slightly involved due to the conditioning on the identity. An alternative approach, perhaps surprisingly, is to abandon the model in favor of one that is more complicated (and realistic). The next section describes such a model. Among other things, the incorporation of discreteness introduces a freedom of motion for c_t that simplifies the sampling.

4. Discreteness.

In the original Roll model, bids, asks and transaction prices are considered to be continuous random variables. In fact, virtually all markets constrain the support of these quotes to a discrete lattice defined as integer multiples of the "tick" or "pip". The tick size is of economic interest because it is related to the cost of achieving time priority, and therefore to the supply of liquidity (Harris (1997a); Harris (1997b)). From a data-modeling perspective, the tick size is important because it is often of magnitude similar to that of the spread and short-term price movements. Harris (1990a) suggests a latent-variable model of rounded transaction prices. Hasbrouck (1998a) surveys this and other approaches taken to modeling discreteness, and proposes the model used below.

a. The model and associated Gibbs sampler

Taking the previous model as a starting point, equation (2) is modified to reflect a rounding transformation:

$$\begin{aligned} B_t &= \text{Floor}[M_t - C_t] \\ A_t &= \text{Ceiling}[M_t + C_t] \end{aligned} \tag{9}$$

where B and A are the level bid and ask and $M_t = \exp(m_t)$. $\text{Floor}[\cdot]$ and $\text{Ceiling}[\cdot]$ round their arguments asymmetrically, down and up (respectively) to the next grid point. (It is assumed that the data are scaled so that the tick size is unity.) The price dynamics for the implicit efficient price are the same as in the previous model, the log random walk given in equation (1). Quote discreteness in the model is (as in reality) imposed on the levels. The cost variable C_t is now stated in level terms. It is considered to be a nonnegative random variable. From an economic perspective, C_t may most conveniently be interpreted as the marginal cost of market-making. The asymmetric rounding ensures that the dealer faces no expected loss. Hasbrouck discusses further economic aspects of this model.

Conditional parameter draws

Economic theory gives little guidance in choice of distribution for C_t, other than the presumption that it must be nonnegative.[9] Two obvious candidate distributions are the lognormal (used in Hasbrouck (1998b)) and exponential. Both are easy to draw from, and both are easily parameterized within a Bayesian framework. As in the previous model, the inverted gamma distribution is a convenient prior for σ_u^2. The update and

[9] While the "dealer cost" interpretation of C_t suggested above makes nonnegativity appear reasonable. Broader interpretations may render it questionable (see Hasbrouck (1998a)).

posterior draw are slightly different, however. The u_t employed for the update are no longer regression residuals (as in equation (5)). They are instead computed from the simulated m_t: $u_t = m_t - m_{t-1}$.

Conditional latent data draw

The procedure employed in the last section must be modified as follow. At each point in time, the model now possesses three unobserved (latent) state variables: m_t (or equivalently M_t), q_t and C_t. It might be supposed that simulation of these three variables could be achieved by a succession of Gibbs draws. This turns out to be only partially true, however.

The cost parameter C_t may indeed be drawn from its full conditional distribution $f(C_t | M_t, q_t, P_t)$. (For notational simplicity, the parameters of the unconditional distribution of C_t have been suppressed.) Given M_t and q_t we know whether the trade price was the bid or the ask. From equation (9), this imposes truncation bounds on C_t:

$$\begin{aligned} &P_t - 1 - M_t < C_t < P_t - M_t, \text{ if } q_t = +1 \\ &M_t - P_t - 1 < C_t < M_t - P_t, \text{ if } q = -1 \end{aligned} \tag{10}$$

We may simply draw C_t from its unconditional distribution, subject to these bounds. These bounds give C_t a latitude of motion, in contrast to the random-cost model suggested in Section 3 prior to our consideration of discreteness.

Matters are not so simple for the other two variables. In the model of Section 3, knowledge of p_t, c and q_t suffices to determine m_t. This is no longer true. Even if we condition on C_t, the discreteness transformation only serves to bound M_t. Suppose that we attempt to construct a Gibbs sampler along the following lines. We will first draw M_t from $f(M_t | M_{t-1}, M_{t+1}, P_t, q_t, C_t)$ and then draw q_t from $f(q_t | M_{t-1}, M_t, M_{t+1}, P_t, C_t)$. Suppose that going into the first M_t draw we have q_t=+1 ("a buy order"). This means that the transaction price is equal to the implicit ask quote, which implies (from the bounds in (9)) that $M_t < P_t - C_t$. When we move to the q_t draw, this last restriction implies q_t=+1.

For an MCMC simulator to work, it must be capable of reaching all points in the support of the density. In this example, the Gibbs sampler fails because q_t=–1 can never be realized.

Fortunately it is easy to draw from the joint bivariate conditional density $f\left(m_t, q_t \middle| m_{t-1}, m_{t+1}, p_t, C_t\right)$. First note that $f\left(m_t \middle| m_{t-1}, m_{t+1}\right)$ is unchanged from equation (6) Given C_t equation (9) imposes truncation bounds. In terms of the level variable M_t:

$$\begin{aligned} &P_t - C_t - 1 < M_t < P_t - C_t, \text{if } q_t = +1 \\ &P_t + C_t < M_t < P_t + C_t + 1, \text{if } q_t = -1 \end{aligned} \tag{11}$$

For notational convenience, denote the set of m_t that satisfies the appropriate restriction as $P_t^{-1}(C_t, q_t)$. The conditional probability that P_t is at the ask is:

$$\Pr\left[q_t \middle| m_{t-1}, m_{t+1}, P_t, C_t\right] \propto \int_{m_t \in P_t^{-1}(C_t, q_t)} f\left(m_t \middle| m_{t-1}, m_{t+1}\right) dm_t \tag{12}$$

We only need to compute this for q_t=1±, normalize, and draw q_t from the implied Bernoulli distribution. Next, note that $f\left(m_t \middle| m_{t-1}, m_{t+1}, P_t, q_t\right)$ is proportional to $f\left(m_t \middle| m_{t-1}, m_{t+1}\right)$ truncated to the region implied by q_t in equation (11), so we may simply make the truncated draw.[10]

b. Application to the futures data

The discreteness model was implemented for the three representative contracts for cost distributions assumed to be either lognormal or exponential Over a wide range of priors and starting values, however, the convergence properties of the estimators was poor. The results for the exponential-cost model of the pork belly contract (Figure 2) are

[10] It might be supposed that a simpler model would result by forcing C_t to be constant over time (as in the original Roll model). Although this simplifies the model in a conceptual sense, it leads to a major degradation of the Gibbs sampler outlined above. If $C_t = C$ is regarded as fixed over time, a new value of C must be drawn subject to the *intersection* of the bounds given in (10) for all t. This severely restricts the extent to which C can change in successive draws. The problem is aggravated in large samples. Simulations confirmed that the sampled values of C exhibited extremely large persistence.

typical. The two model parameters are the mean of the exponential distribution ("Mu_C") and σ_u ("SD_u"). The draws (particularly for the cost parameter) manifest poor mixing and large persistent deviations. Autocorrelations are high.

In fixing the cause of this disappointing performance, suspicion must first fall on the estimation methodology. The Gibbs sampler used here is a single-step procedure. It is known that when parameters and/or latent data are highly correlated, such samplers are prone to poor convergence, and should be avoided in favor of "block" samplers. The convergence problem did not generally arise, however, in simulated data sets. Thus, although this possibility cannot be excluded entirely, it does not seem to be arising from the limitations of the sampler.

It is more reasonable to conjecture that model misspecification and data limitations preclude reliable identification of the cost parameters. Even if we could observe the actual bid and ask quotes, inference about the cost parameters would be based on grouped data (where the grouping is driven by discreteness). If the grouping is coarse (as seems likely) in the present application, the inference is severely impaired. Matters are further aggravated here because the bid and ask are not observed. Furthermore, in studies where the bid and ask are observed, the lognormal distribution appears to imply large-spread occurrence frequencies that are higher than those found in the data. (Manrique and Shephard (1997)) note this for an NYSE stock; Hasbrouck (1998b), for Deutschemark/US dollar quotes in the interbank FX market.)

Correct modeling of implicit quote exposure costs thus remains as an important area of further research. For present purposes, however, the apparent lack of identification of this cost will be handled by suppressing this cost entirely, that is, by setting C_t=0. This does not, of course, force the spread to zero: the rounding mechanism specified in (9) ensures a one-tick spread. Spreads larger than this will be assumed generated by clustering (as described in the next section), and we defer further estimations to then.

From a methodological viewpoint, the decision to set the implicit continuous cost to zero does not arise from a belief that these costs are economically zero. It is, rather, a frank admission that these costs cannot be reliably estimated. In models that ignore discreteness and clustering, apparently well-behaved estimates will arise (as in the previous section). These estimates are, however, artifacts of discreteness and clustering transformations that are more properly modeled directly.

5. Clustering

As noted in Section 2, futures transaction prices frequently exhibit pronounced clustering. Hasbrouck (1998b) suggests that clustering in bid and ask quotes be modeled as a consequence of an implicit tick, a natural multiple of the minimum tick, that arises as a trading convention or from individual preference. Clustering in bid and ask quotes naturally gives rise to clustering in the transactions that occur at these quotes. The results presented in Table 4 suggest that pork belly prices exhibit a strong tendency to cluster on even ("two-multiple") prices. S&P contract prices have a modest preference for multiples of five. The Deutschemark contract prices are not strongly clustered.

a. The model and associated Gibbs sampler

Clustering is imposed on the (unobserved) bid and ask quotes by using generalized rounding functions:

$$\begin{aligned} B_t &= \text{Floor}[M_t - C_t, K_t] \\ A_t &= \text{Ceiling}[M_t + C_t, K_t] \end{aligned} \tag{13}$$

where K_t denotes the tick-multiple to which rounding will occur. In economic terms, K_t is the implicit tick size. While K_t might modeled in a very general fashion, the specifications estimated here will allow for only two possible values: one (that is, the regular tick increment) and κ, a single dominant multiple. In the present analyses, κ=2 for the pork belly contract and κ=5 for the S&P contract. Although the Deutschemark contract prices are not strongly clustered, for the sake of estimating all specifications in parallel, clustering with κ=2 will be allowed. As in Hasbrouck (1998b), it is convenient to assume an i.i.d. Bernoulli distribution:

$$K_t = \begin{cases} 1, \text{ w. prob. } (1-k) \\ \kappa, \text{ w. prob. } k \end{cases} \tag{14}$$

The Bernoulli probability parameter k may be interpreted as the clustering intensity.

Hasbrouck (1998b) uses this framework in modeling foreign exchange bids and asks that are actually observed. The present situation is more challenging in that we do not observe these quotes directly, only transaction prices. The Gibbs procedure proposed in the last section must be modified in two respects. There is, firstly, one more parameter to be estimated (k) in the parameter draw step. Secondly, the latent data draw at each time t now involves four latent state variables: M_t, C_t, q_t and K_t.

Conditional parameter draw

The parameter draw is uncomplicated. The number of "hits" (instances of K_t=κ) in a sample of T observations is a binomial random variable. A conjugate prior for the hit probability is the Beta distribution. The posterior is updated on the basis of the (simulated) K_t, and a random value of k is drawn. (See the appendix.)

Conditional data draw

The latent data draw is slightly more involved. The draw for C_t is subject to truncation bounds derived from (13) (cf. equation (10) for the nonclustered case):

$$\begin{aligned} P_t - K_t - M_t < C_t < P_t - M_t, \text{ if } q_t = +1 \\ M_t - P_t - K_t < C_t < M_t - P_t, \text{ if } q_t = -1 \end{aligned} \tag{15}$$

As before, we may draw from the unconditional distribution of C_t subject to the relevant truncation bound.

As in the previous model, M_t (or m_t) and q_t must be drawn jointly. It is also convenient (and computationally efficient) to draw K_t jointly as well. Given P_t, K_t and q_t, the bounds on M_t implied by (13) are:

$$P_t - C_t - K_t < M_t < P_t - C_t, \text{ if } q_t = +1$$
$$P_t + C_t < M_t < P_t + C_t + K_t, \text{ if } q_t = -1 \qquad (16)$$

As above, denote the set of m_t that satisfies this restriction as $P_t^{-1}(K_t, C_t, q_t)$.

If the transaction price P_t does *not* lie on a κ-multiple, we can't have clustering. In this case K_t=1 and we may make the joint draw of M_t and q_t exactly as in the previous model. On the other hand, the observation that P_t *does* lie on a κ-multiple does not imply K_t=κ. (It might be the case that K_t=1 and the configuration of the other variables maps to a transaction price that just happens to be a κ-multiple.)

In drawing q_t and K_t here there are four possibilities: $(q_t, K_t) \in \{q_t = +1, q_t = -1\} \times \{K_t = 1, K_t = \kappa\}$. Although q_t and K_t are unconditionally independent (by assumption), they are dependent conditional on other model variables. The joint conditional distribution of q_t and K_t therefore has probabilities:

$$\begin{aligned} \Pr[q_t, K_t \mid P_t, C_t, m_{t-1}, m_{t+1}] &\propto \Pr[q_t]\Pr[K_t]\Pr\left[m_t \in P_t^{-1}(K_t, C_t, q_t)\right] \\ &= \Pr[q_t]\Pr[K_t]\int_{m_t \in P_t^{-1}(K_t, C_t, q_t)} f\left(m_t \middle| m_{t-1}, m_{t+1}\right) dm_t \end{aligned} \qquad (17)$$

where $\Pr[q_t]$ and $\Pr[K_t]$ are the unconditional probabilities and the conditional density for m_t is given in equation (6). To proceed, we compute the r.h.s. of (17) for all values of q_t and K_t, normalize to obtain the l.h.s. conditional probabilities, and make the joint draw of q_t and K_t. Finally, given q_t and K_t we draw m_t from the conditional distribution $f\left(m_t \middle| m_{t-1}, m_{t+1}\right)$ subject to the truncation $m_t \in P_t^{-1}(K_t, C_t, q_t)$.

b. Application to the futures data

The model with discreteness and clustering was estimated for the three representative contracts. Following the remarks in the last section, the quote exposure cost C_t was fixed at zero, thus forcing discreteness and clustering to account for all of the market's bid-ask spread. Visual monitoring of convergence suggested that that the Gibbs sampler performed well.

Table 8 summarizes the results. Most importantly, estimates of the clustering probability parameters (the *k*'s) are consistent with the tick-multiple statistics in Table 4.

For the pork belly contract, k=74%. The S&P contract exhibits low, but discernible clustering. For the Deutschemark contract, k is virtually indistinguishable from zero. For all three contracts, the estimated sign classification accuracy (π^C) is high. In the case of the pork belly contract, this represents a significant improvement from the value associated with the model absent discreteness and clustering (cf. Table 7).

6. Asymmetric information

The basic model and the variants presented above assume that the innovation to the efficient price is independent of the incoming order, i.e., that the quote setter infers nothing from this order. This is particularly restrictive given the modern view that a security market should function as an aggregator of diverse private information. An essential characteristic of the sequential trade models is the possibility that the incoming trade is a signal for the traders private information, and that the quote setter will make optimal use of this signal in updating her bid and ask.

The introduction of asymmetric information complicates the model's conditional distributional in certain respects. Although the full model eventually estimated will allow for the imperfections discussed in earlier sections, it is best for expositional purposes to examine asymmetric information apart from discrete, clustering and random costs of quote exposure. Accordingly, the simplified model discussed below is a straightforward modification of the basic Roll model discussed in Section 3.

A modification of equation (1) that permits the incoming trade to affect the efficient price evolution is

$$m_t = m_{t-1} + \lambda q_t v_t + u_t \tag{18}$$

where v_t is the unsigned transaction volume (e.g., 100 shares, contracts or whatever), q_t is (as before) the direction of the trade, and λ is an impact coefficient (sometimes termed the "liquidity" parameter). This interpretation of the $\lambda q_t v_t$ term is intuitively useful, but the following developments are considerably more general. If we only know that a trade has occurred, but are ignorant of its size, the term can be replaced by "λq_t", in which case λ will reflect the directional impact of a trade of unknown size. Alternatively, v_t may be a vector-valued transformation of the trade size (possibly including an intercept, linear and quadratic terms as in Hasbrouck (1991a)), in which case λ is a coefficient vector. Although this formulation presumes that buys and sells affect the price in a symmetric fashion, this too could be generalized.

From an economic perspective, $\lambda q_t v_t$ reflects the price adjustment based on the signal of private information that the quote-setter infers from the trade. The disturbance u_t plays a narrower role than before. It now solely reflects non-trade public information, and is assumed independent of q_t and v_t. It is provisionally assumed that trades take place exactly at the efficient prices, i.e., $p_t = m_t$. This is along the lines of the Glosten and Milgrom (1985) model, with no trading costs (aside from those associated with the informational asymmetry).

Equation (18) may be written as the regression:

$$\Delta p_t = \Delta m_t = \lambda q_t v_t + u_t \tag{19}$$

This implies a variance decomposition for the log efficient price changes:

$$\begin{aligned} Var(\Delta m_t) &= E(q_t \lambda v_t)^2 + \sigma_u^2 \\ \underbrace{\qquad\qquad}_{\text{Total}} &= \underbrace{E(\lambda v_t)^2}_{\substack{\text{Trade-related} \\ \text{contribution}}} + \underbrace{\sigma_u^2}_{\substack{\text{Public (non-trade-related)} \\ \text{contribution}}} \end{aligned} \tag{20}$$

This decomposition has economic content in that it highlights the relative importance of trading for the price discovery process (Hasbrouck (1991b)).

Estimation is based on a data record of prices and volumes for each trade: p_t and v_t for $t = 1, \ldots, T$. There are two model parameters, λ and σ_u^2; the latent variables are q_t, $t = 1, \ldots, T$.

Conditional parameter draw

The conditional parameter draws (assuming that the q_t are known) are straightforward and similar to those for the original Roll model. We simply estimate equation (19) in a Bayesian regression framework, subject to the identification restriction $\lambda \geq 0$.

Conditional data draw

Conditional simulation of the latent trade direction q_t proceeds as follows. First note that the conditional distribution of m_t is:

$$f\left(m_1 | q_1, m_2, q_2\right) = f\left(m_1 | m_2, q_2\right) = N\left(m_2 - \lambda q_2 v_2,\ \sigma_u^2\right)$$
$$f\left(m_t | m_{t-1}, q_t, q_{t+1}, m_{t+1}\right) = N\left(\frac{\left(m_{t-1} + m_{t+1} + \lambda q_t v_t - \lambda q_{t+1} v_{t+1}\right)}{2}, \frac{\sigma_u^2}{2}\right) \text{ for } t = 2, \ldots, T-1 \qquad (21)$$
$$f\left(m_T | m_{T-1}, q_T\right) = N\left(m_{T-1} + \lambda q_T v_T, \sigma_u^2\right)$$

In contrast to all of the earlier models the conditional density for m_t generally depends (via the mean) on q_t.

The conditional probability of the trade sign is:

$$\Pr\left[q_t | q_{/t}, p\right] = \Pr\left[q_t | m_{t-1}, m_{t+1}, q_{t+1}, p_t\right] \propto f\left(m_t | m_{t-1}, q_t, m_{t+1}, q_{t+1}\right) \Pr\left[q_t | m_{t-1}, q_{t+1}, m_{t+1}\right] \qquad (22)$$

evaluated at $m_t = p_t$. In contrast to the earlier analyses, however, $\Pr\left[q_t | m_{t-1}, q_{t+1}, m_{t+1}\right] \neq \Pr\left[q_t\right] = 1/2$, so a further computation is required.

$$\Pr\left[q_t | m_{t-1}, q_{t+1}, m_{t+1}\right] \propto \int_{-\infty}^{+\infty} f\left(m_t, m_{t+1} | m_{t-1}, q_t, q_{t+1}\right) dm_t \qquad (23)$$

where $f\left(m_t, m_{t+1} | m_{t-1}, q_t, q_{t+1}\right) = f\left(m_{t+1} | m_t, q_{t+1}\right) f\left(m_t | m_{t-1}, q_t\right)$, given the structure of the model.

We proceed by first computing the r.h.s. of (23) for $q_t = +1$ and $q_t = -1$. (The integration is trivial given the normality of u_t.) We normalize to obtain the l.h.s. probabilities. We plug the latter into the r.h.s. of (22) (for $q_t = +1$ and $q_t = -1$) and normalize (again) to obtain the l.h.s. probabilities. The latter are the conditional

Bernoulli probabilities necessary to make the draw of q_t. Given q_t, we draw m_t from $f\left(m_t \middle| m_{t-1}, q_t, q_{t+1}, m_{t+1}\right)$ or its variant given in equation (21).

When we incorporate discrete, clustering and random costs of market making, the first set of computations (those related to (23)) are unaffected. The conditional density for m_t on the r.h.s. of (22), however, is replaced by the integration of this density over the relevant feasible region for m_t.

7. A generalized asymmetric information model

The model presented and estimated in this section combines features discussed above. The log efficient price dynamics are:

$$m_t = m_{t-1} + \lambda_0 q_t v_t + \lambda_1 q_{t-1} v_{t-1} + u_t \text{ where } u_t \sim N\left(0, \sigma_u^2\right)$$

where q_t is the trade direction indicator. In this model, v_t is a bivariate function of the trade volume: $v_t = \begin{bmatrix} 1 & \sqrt{Volume_t} \end{bmatrix}'$. The square-root transformation motivated by trade-price impact studies in equity markets that generally find concavity in the relation. An intercept is included to allow for non-size-related directional effects. The λ's are conformable coefficient vectors: $\lambda_i = \begin{bmatrix} \lambda_{i,Const} & \lambda_{i,\sqrt{}} \end{bmatrix}$ for i=0,1. Identification is ensured by requiring all elements of the λ's to be nonnegative.

The specification permits dependence on lagged trades, effectively allowing the impact of a trade to be distributed over time. The theoretical sequential trade models generally assume that adjustment is instantaneous, but delays in information transmission (even across a trading pit) and trader reaction times may quite plausibly result in lagged adjustment. Lags have also been found useful in the corresponding equity market specifications.

The remaining elements of the model define the clustering:

$$K_t = \begin{cases} 1, \text{ w. prob. } (1-k) \\ \kappa, \text{ w. prob. } k \end{cases}$$

quote formation:

$$B_t = \text{Floor}\left[M_t, K_t\right]$$
$$A_t = \text{Ceiling}\left[M_t, K_t\right]$$

where $M_t = \exp(m_t)$, and transaction price formation:

$$P_t = \begin{cases} B_t \text{ if } q_t = -1 \\ A_t \text{ if } q_t = +1 \end{cases}$$

As in the implementations of the previous models with discreteness and clustering, the rounding is assumed based directly on M_t, with no additional implicit cost of market making.

The Gibbs samplers for this model for the three representative contracts were generally well-behaved. Table 9 summarizes the coefficient posteriors (based on diffuse priors). The clustering probabilities *k* and direction classification probabilities π^C are similar to those from the simpler model of discreteness and clustering found earlier.

Are the trade impact coefficients statistically significantly different from zero? Since they are estimated from priors with nonnegative support, the answer is, "Yes, by assumption". It is perhaps more useful to gauge the implied economic significance of these estimates. We offer two approaches

First, Figure 3 graphs the price impact functions implied by the estimates. The vertical scale (change in log price times one hundred) is approximately "percentage price change" associated with a purchase of a given number of contracts. Each graph reports two curves, corresponding to the lag zero and (cumulative) lag one impacts. For a trade of, say, ten contracts, the price impact is largest for the pork belly contract (approximately 11 basis points), lower for the Deutschemark contract (0.7 basis points) and lowest for the S&P contract (0.3 basis points).

Secondly, it was suggested in Section 6 that (by way of comparison to the equity studies) it is useful to compare the magnitudes of trade- and non-trade-related sources of volatility. Table 10 summarizes these calculations. In accounting for total volatility, trades appear most important for the Deutschemark contract (roughly forty percent), of lesser importance for the pork belly contract (twenty percent), and markedly less important for the S&P contract (three percent). A corresponding figure for an NYSE equity might be around twenty percent (Hasbrouck (1991b)), i.e., close to the middle of the three futures contracts considered here.

The economic models of sequential trade identify permanent trade price impacts with asymmetric information, private information that can be revealed in the price only through trade. From this perspective, it is perhaps not surprising that a significant proportion of volatility in the pork belly market originates from trades. There are, for this contract, few alternative sources of price discovery.

For the Deutschemark, on the other hand, the volume of trade that occurs in the futures market is small relative to that occurring in the interbank spot and forward markets.[11] The conventional view is that price discovery occurs in the interbank market. The futures contract is often supposed to serve as a hedging and speculation vehicle for participants too small to obtain easy access to the larger market. One would therefore expect DM futures prices to follow passively the path established in the interbank market. The contribution of futures trades to price discovery implied by the present model (forty percent) seems implausibly high.

To explore this further, we consider absolute trade impacts. One hundred DM contracts in the futures market, a fairly substantial trade, corresponds to DM 12.5 million (roughly $7 million). By the scale of the interbank market, however, this is a very modest trade. The permanent price impact implied by the present estimates is 0.00014, i.e., 1.4 basis points. The interbank market conventionally quotes in DM per U.S. dollar (the reverse of the futures convention). At the sample average price, this corresponds to roughly 1.8 DM/$. At this level, the 1.4 basis point shift would imply a movement of

---

[11] Recent microstructure studies of the latter include Lyons (1995); Goodhart, Ito, and Payne (1996); Lyons (1997) and Evans (1998).

0.000025. This large, but not grossly out of line with trade/price movements found on the electronic limit order books systems (EBS and Reuters D2000-2).

An important consideration here is that transparency in the interbank market is low. The public record of the interbank market is limited to indicative (nonfirm) bids and offers. Trades that occur on the electronic book systems are visible only to other subscribers, a essentially the large intermarket banks themselves. Neither trades occurring directly between two participants nor those mediated by brokers are publicly reported.

In the interbank market, the tick (pip) size is 0.0001 DM/$. On a relative value basis, this is about half the 0.0001 $/DM tick for the futures contract. Hasbrouck (1998b) nevertheless finds that in 1996, the average spread is roughly six ticks (pips), and the quotes are highly clustered at five-tick multiples. Thus, the implied price impact of the futures trade estimated above falls well within the typical spread. The usefulness of the Reuters indicative quotes as a timely, high-resolution signal for futures price discovery appears doubtful. It seems reasonable to hypothesize that trades in the futures market are driven by information that may well have originated in the interbank market, but is "private" in the sense of not being widely reported. Far from being a subsidiary player in this market, the futures market may be serving as the primary public forum of price discovery.

We turn now to discussion of the trade impacts for the S&P contract. In absolute terms (Figure 3) and especially in relative terms (Table 1) these are extremely, perhaps implausibly, low. While the cash market exists as a meaningful alternative for price discovery, the stock index futures market is customarily viewed as originating the primary signals of common factor equity movements. Both the numerous studies documenting index price leadership in the futures markets, and the studies that address regulatory concerns support this view.

In considering model adequacy, it is noteworthy that the index futures market is substantially more active than the other two contracts. Section 2 noted an average intertrade time of four seconds and raised the possibility of associated informational delays. It is highly likely that for this contract, the one lag allowed in the model for the trade impact is much too abbreviated. This possibility bears further investigation.

8. Conclusions

This paper proposes and implements powerful strategies to estimate empirical microstructure models in the absence of a full data record. The centerpiece model is a structural model of bid and ask quotes and trades that incorporates discreteness, clustering and asymmetric information. Yet for all its richness, it can be estimated solely from reported transaction (price and volume) data. The analysis is made possible by recent advances in Markov chain Monte Carlo estimation, which simplify inference in latent-data models.

The paper presents a preliminary analysis of a dozen high-volume contracts, and a more detailed study of three representative contracts: pork bellies, the Deutschemark and the S&P Composite Index.

Preliminary analysis of futures transaction price data suggests price clustering (affinity for natural multiples of the minimum tick) that is, for certain contracts, quite pronounced. Clustering is very strong for the pork belly contract; small for the S&P contract and negligible for the DM contract. It is not determined whether this clustering arises from negotiation-cost minimization or market power of floor traders.

The model also provides evidence on trade-price impacts and the importance of trades as sources of permanent price movements. For the S&P contract, the price impact of a trade is low: a hypothetical ten-contract purchase order moves the price by only 0.3 basis points (0.00003%). Of the total permanent volatility, only three percent is attributed to trades. Taken at face value, this implies a minimal informational role for the

S&P index futures market. The estimated specification, however, permits only the current and most recent lagged trade to drive the price change. The pace of trading in the S&P contract is sufficiently high that price impacts may be distributed over a greater number of lagged trades.

For the pork belly contract, the estimated impact of a 10-contract purchase is 11 basis points, and roughly twenty percent of the long-term price volatility is attributed to trades. The latter figure is comparable to that found in equity market studies and suggests a strong informational role for trading.

The estimated impact of ten-contract purchase in the DM contract is low (0.7 basis points), but the share of long-term volatility attributed to trades is, at forty percent, the highest. The latter figure implies that futures market trading contributes significantly to the price discovery process. This runs counter to the conventional wisdom that price determination in foreign exchange occurs in the interbank spot/forward market. Transparency in the interbank market, however, is low. Given that interbank trades are not reported, it is perhaps not surprising that the publicly-reported (though smaller) futures trades play a substantial role in price discovery.

Appendix: Standard Bayesian Results

For the reader's convenience, this appendix summarizes some key Bayesian results used in the body of the paper. They are fairly standard (see, for example, Carlin and Louis (1996), Tanner (1996), Press (1989) or Zellner (1971)).

a. Univariate normal random variables.

Suppose that a random variable $x \sim N(\mu, \sigma^2)$, and that we possess a sample of independent observations $x_i, i = 1, \ldots, n$. A convenient (conjugate) prior for the mean parameter is $\pi(\mu) = N(\mu^{prior}, \tau^2)$. Suppose that the variance σ^2 is known. Then the posterior is $f(\mu|x) = f(\mu|\bar{x}) = N(\mu^{post}, \tau^{2,post})$, where

$$\mu^{post} = \frac{\sigma^2 \mu^{prior} + n\tau^2 \bar{x}}{\sigma^2 + n\tau^2};\ \tau^{2,post} = \frac{\sigma^2 \tau^2}{\sigma^2 + n\tau^2} \tag{24}$$

In the limit as $\tau \to \infty$, we arrive at the uninformative (improper) prior with corresponding posterior parameters $\mu^{post} = \bar{x};\ \tau^{2,post} = \sigma^2/n$. In any case, a random posterior draw consists of a draw from a normal density.

Some model parameters in this paper have truncated priors. (Most commonly, a parameter is asserted to be nonnegative.) Denote by $N_T(\mu, \tau^2)$ the normal density with mean μ and variance τ^2 truncated to some region of the real line. The truncation changes the normalization of the density. Furthermore if $x \sim N_T(\mu, \tau^2)$, $Ex \neq \mu$ and $\text{Var}(x) \neq \tau^2$. Nevertheless, the above results go through as before. That is, if our prior is $\pi(\mu) = N_T(\mu^{prior}, \tau^2)$, then the posterior is $f(\mu|x) = f(\mu|\bar{x}) = N_T(\mu^{post}, \tau^{2,post})$, i.e., a density with the same posterior parameters and a truncation region that is identical to the truncation of the prior. NB: this result goes through because the truncation region is not data-dependent.

With normal observations, a convenient prior for the variance parameter is $\pi(\sigma^2) = IG(\alpha^{prior}, \beta^{prior})$, the inverse gamma distribution. Based on a sample of observations iid normal with known mean μ, the posterior is $f(\sigma^2|x) = IG(\alpha^{post}, \beta^{post})$ where

$$\alpha^{post} = \alpha^{prior} + \frac{n}{2};\ \beta^{post} = \left[\frac{1}{\beta^{prior}} + \frac{\sum (x_i - \mu)^2}{2}\right]^{-1} \tag{25}$$

The improper prior $\pi(\sigma^2) \propto 1/\sigma^2$ is obtained by letting $\alpha^{prior} = \beta^{prior} = 0$. The posterior is then proportional to the inverse chi-square density: $\sigma^2 | x \sim \sum (x_i - \mu)^2 \chi^{-2}_{df=n}$

b. The Bayesian Gaussian linear model

The model is $y_i = x_i\beta + u_i$ where x_i is a row vector of known data, β is a column coefficient vector and $u_i \overset{iid}{\sim} N(0, \sigma_u^2)$. A conjugate prior for the coefficient vector is the multivariate normal: $\pi(\beta) = N(\mu^{prior}, \Sigma^{prior})$. Assuming σ_u^2 known, the posterior is $f(\beta|y) = N(\mu^{post}, \Sigma^{post})$ where

$$\begin{aligned}(\Sigma^{post})^{-1} &= \frac{(X'X)}{\sigma_u^2} + (\Sigma^{prior})^{-1} \\ \mu^{post} &= \Sigma^{post}\left[\frac{X'y}{\sigma_u^2} + (\Sigma^{prior})^{-1}\mu^{prior}\right]\end{aligned} \tag{26}$$

The improper coefficient prior is obtained by letting Σ^{prior} become "large" (while remaining positive definite). If the prior is truncated to some region, the posterior is also truncated. Assuming β known, the variance parameter σ_u^2 may be handled exactly as in the normal univariate case above, using the model residuals $u_i = y_i - x_i\beta$ in lieu of the deviations from the mean.

c. The Bernoulli/Binomial Model

Suppose x_i is a Bernoulli random variable:

$$x_i = \begin{cases} 1, \text{ w. prob. } k \\ 0, \text{ w. prob. } (1-k) \end{cases} \tag{27}$$

A convenience prior for the probability parameter *k* is the beta distribution, denoted $B(\alpha^{prior}, \beta^{prior})$. Setting $\alpha^{prior} = \beta^{prior} = 1$ gives the uniform prior; the Jeffreys (noninformative) prior is obtained with $\alpha^{prior} = \beta^{prior} = 1/2$. Suppose that we possess a sample of *N* observations, of which *n* are "hits" (instances of x_i=1). The posterior for *k* is $B(\alpha^{post}, \beta^{post})$ where $\alpha^{post} = \alpha^{prior} + n$ and $\beta^{post} = \beta^{prior} + N - n$.

References

Carlin, B. P., and T. A. Louis. 1996. *Bayes and Empirical Bayes Methods for Data Analysis*. London: Chapman and Hall.

Carlin, B. P., N. G. Polson, and D. S. Stoffer. 1992. A Monte Carlo Approach to Nonnormal and Nonlinear State Space Modeling. *Journal of the American Statistical Association* 87, no. 418: 493-500.

Carter, C. K., and R. Kohn, 1994, On Gibbs Sampling for State Space Models, *Biometrika*, 81, 541-53.

Casella, G., and E. I. George, 1992, Explaining the Gibbs Sampler, *The American Statistician*, 46, 167-90.

Chib, S., and E. Greenberg. 1996. Markov Chain Monte Carlo Simulation Methods in Econometrics. *Econometric Theory* 12: 409-31.

Christie, W. G., J. H. Harris, and P. H. Schultz, 1994, Why did NASDAQ market makers stop avoiding odd-eighth quotes?, *Journal of Finance*, 49, 1841-60.

Dutta, P. K., and A. M. Madhavan, 1997, Competition and Collusion in Dealer Markets, *Journal of Finance*, 52, 245-76.

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

———. 1991. Order Form and Information in Securities Markets. *Journal of Finance* 46, no. 3: 905-27.

———. 1992a. Adverse Selection and Large Trade Volume: The Implications for Market Efficiency. *Journal of Financial and Quantitative Analysis* 27, no. 2: 185-208.

———. 1992b. Time and the Process of Security Price Adjustment. *Journal of Finance* 47, no. 2: 576-605.

Evans, M., 1998, The microstructure of the foreign exchange market, working paper, Department of Economics, Georgetown University.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996. Introducing Markov Chain Monte Carlo. *Markov Chain Monte Carlo in Practice.* Editors W. R. Gilks, S. Richardson, and D. J. Spiegelhalter London: Chapman and Hall.

Glosten, L. R., and L. E. Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics,* 21, 123-42.

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

Goodhart, C., T. Ito, and R. Payne. 1996. One Day in June 1993: A Study of the Working of the Reuters 2000-2 Electronic Foreign Exchange Trading System. *The Microstructure of Foreign Exchange Markets.* Editors J. A. Frankel, G. Galli, and A. Giovannini Chicago: University of Chicago Press.

Harris, L. E., 1990a, Estimation of stock price variances and serial covariances from discrete observations, *Journal of Financial and Quantitative Analysis,* 25, 291-306.

———, 1990b, Statistical properties of the roll serial covariance bid/ask spread estimator, *Journal of Finance,* 45, 579-90.

———, 1991, Stock price clustering and discreteness, *Review of Financial Studies,* 4, 389-415.

———, 1994, Minimum price variations, discrete bid-ask spreads, and quotation sizes, *Review of Financial Studies,* 7, 149-78.

———, 1997a, Decimalization: a review of the arguments and evidence, working paper, University of Southern California.

———, 1997b, Does a large minimum price variation encourage order exposure?, working paper, School of Business Administration, University of Southern

California.

Hasbrouck, J., 1988, Trades, quotes, inventories, and information, *Journal of Financial Economics,* 22, 229-52.

———, 1991a, Measuring the information content of stock trades, *Journal of Finance,* 46, 179-207.

———. 1991b. The Summary Informativeness of Stock Trades: An Econometric Analysis. *Review of Financial Studies* 4, no. 3: 571-95.

———, 1992, Using the TORQ Database, working paper, New York Stock Exchange.

———, 1996a, Modeling microstructure time series, in G. S. Maddala and C. R. Rao (editors), Handbook of Statistics 14: Statistical Methods in Finance, Amsterdam: Elsevier North Holland.

———. 1996b. Order Characteristics and Stock Price Evolution: An Application to Program Trading. *Journal-of-Financial-Economics* 41, no. 1: 129-49.

———, 1998a, The dynamics of discrete bid and ask quotes, *Journal of Finance*. forthcoming.

———, 1998b, Security bid/ask dynamics with discreteness and clustering: simple strategies for modeling and estimation, working paper, Stern School, New York University, www.stern.nyu.edu\~jhasbrou.

Hasbrouck, J., and T. S. Y. Ho. 1987. Order Arrival, Quote Behavior, and the Return-Generating Process. *Journal of Finance* 42, no. 4: 1035-48.

Jacquier, E., N. G. Polson, and P. E. Rossi, 1994, Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics,* 12, 371-89.

Kandel, E., and L. M. Marx, 1997, NASDAQ Market Structure and Spread Patterns, *Journal of Financial Economics,* 45, 61-89.

Kim, S., N. Shephard, and S. Chib, 1998, Stochastic volatility: likelihood inference and comparison with ARCH models, *Review of Economic Studies,* forthcoming.

Kitagawa, G., 1987, Non-Gaussian state-space modeling of nonstationary time series, *Journal of the American Statistical Association*, 82, 1032-41.

Laux, P. A., and A. J. Jr. Senchack, 1992, Bid-ask spreads in financial futures, *Journal of Futures Markets*, 12, 621-34.

Lee, C. M. C., and M. J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance*, 46, 733-46.

Lyons, R. K., 1995, Tests of microstructural hypotheses in the foreign exchange market, *Journal of Financial Economics*, 39, 321-51.

———. 1997. A Simultaneous Trade Model of the Foreign Exchange Hot Potato. *Journal of International Economics* 42, no. 3-4: 275-98.

Ma, C. K., R. L. Peterson, and R. S. Sears, 1992, Trading noise, adverse selection and intraday bid-ask spreads in futures markets, *Journal of Futures Markets*, 12, 519-38.

Madhavan, A., M. Richardson, and M. Roomans, 1997, Why do security prices change?, *Review of Financial Studies*.

Manaster, S., and S. C. Mann, 1996, Life in the pits: competitive market making and inventory control, *Review of Financial Studies*, 9, 953-75.

Manrique, A. 1997. "Econometric Analysis of Limited Dependent Time Series." Ph.D. Dissertation, Nuffield College, Oxford University.

Manrique, A. and Shephard, N., 1997, Likelihood analysis of a discrete bid/ask price model for a common stock , working paper, Nuffield College, Oxford University.

———, 1998, Likelihood inference for limited dependent processes, *Econometrics Journal*.

O'Hara, Maureen. 1995. *Market Microstructure Theory*. Cambridge, USA: Blackwell Publishers.

Odders-White, E. R., 1997, On the occurrence and consequences of inaccurate trade classification, working paper, University of Wisconsin.

Press, S. J. 1989. *Bayesian Statistics: Principles, Models, and Applications*. New York:

John Wiley.

Roll, R. 1984. A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *Journal of Finance* 39, no. 4: 1127-39.

Schwert, G. W., 1997, Symposium on Market Microstructure: Focus on NASDAQ, *Journal of Financial Economics,* 45, 1-8.

Shephard, N. 1993. Fitting Nonlinear Time-Series Models with Applications to Stochastic Variance Models. *Journal of Applied Econometrics* 8, no. Suppl.: S135-S152.

———. 1994. Partial non-Gaussian State Space. *Biometrika* 81, no. 1: 115-31.

Shephard, N., and A. Manrique. 1997. Simulation-Based Likelihood Inference for Limited Dependent Processes. *Working Paper, Nuffield College, Oxford University*.

Shephard, N. and Pitt, M. K., 1997a, Filtering via simulation: auxiliary particle filters, working paper, Nuffield College, Oxford University.

———. 1997b. Likelihood analysis of non-Gaussian Measurement Time Series. *Biometrika* 83: 653-67.

Smith, T., and R. E. Whaley, 1994, Estimating the effective bid/ask spread from time and sales data, *Journal of Futures Markets,* 14, 437-55.

Tanner, M. A. 1996. *Tools for Statistical Inference*. Third ed. New York: Springer-Verlag.

U. S. Commodities Futures Trading Commission. 1997. "Rule Enforcement Review of Exchange Broker Association Programs." Web page. Available at www.cftc.gov.

Zellner, A. 1971. *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley.

Table 1. Contract Descriptions.

Notes: Contracts traded on the Chicago Mercantile Exchange for the indicated underlying and maturity. Averages are computed over all trades from January 5 to January 16, 1998. The units of the average contract value for the Euroyen contract are Y1,000.

Underlying	Maturity, 1998	Average Price	Units	Tick Size	Contract Size	Average Contract Value ($1,000)	Tick Size as % of Average Price
Feeder Cattle	Mar	76.17	Cents/Lb	0.025	50,000 Lb	38	0.033
Live Cattle	Feb	64.64	Cents/Lb	0.025	40,000 Lb	26	0.039
Pork Bellies	Feb	49.48	Cents/Lb	0.025	40,000 Lb	20	0.051
Australian Dollar	Mar	0.65	US$/AD	0.0001	100,000 AD	65	0.015
Canadian Dollar	Mar	0.70	US$/CD	0.0001	100,000 CD	70	0.014
Deutsche Mark	Mar	0.55	US$/DM	0.0001	125,000 DM	69	0.018
Japanese Yen	Mar	0.77	$.01US/JY	0.0001	12.5 MillionY	96	0.013
Swiss Franc	Mar	0.68	US$/SF	0.0001	125,000 SF	85	0.015
13wk Treasury Bill	Mar	95.22	Pts of 100%	0.005	$1 Million	952	0.005
Eurodollar	Mar	94.42	Pts of 100%	0.005	$1 Million	944	0.005
Euroyen	Mar	99.12	Pts of 100%	0.005	100 MillionY	99,124	0.005
One-Month LIBOR	Jan	94.39	Pts of 100%	0.005	$3 Million	2,832	0.005
Nasdaq 100	Mar	1,000.76	Index Pts	0.05	$100 x Index	100	0.005
Nikkei 225 Index	Mar	15,016.08	Index Pts	5	$5 x Index	75	0.033
S&P 500 Index	Mar	961.65	Index Pts	0.1	$250 x Index	240	0.010
S&P Midcap 400 Index	Mar	324.01	Index Pts	0.05	$500 x Index	162	0.015

Table 2. Trading Statistics

Trading activity in the indicated CME contracts from January 5, 1998 to January 16, 1998.

		Contracts per Trade					
	Avg. Trades Per Day	Min	25%'ile	Median	50%'ile	Max	Avg. Inter-trade time (seconds)
Feeder Cattle	301	1	1	3	7	179	47
Live Cattle	604	1	2	5	11	414	23
Pork Bellies	300	1	1	2	4	188	45
Australian Dollar	105	1	1	2	7	296	224
Canadian Dollar	503	1	1	3	10	485	47
Deutsche Mark	895	1	2	6	16	864	26
Japanese Yen	1,211	1	2	5	11	1,550	19
Swiss Franc	908	1	2	4	10	619	26
13wk Treasury Bill	34	1	1	5	25	1,525	686
Eurodollar	492	1	40	100	280	3,587	49
Euroyen	42	1	2	5	27	510	565
One-Month LIBOR	33	1	12	40	80	700	738
Nasdaq 100	484	1	1	4	10	250	50
Nikkei 225 Index	129	1	2	5	10	133	204
S&P 500 Index	5,526	1	3	9	22	805	4
S&P Midcap 400 Index	109	1	2	3	7	52	225

Table 3. Intertransaction Price Properties

Sample of contracts traded on the Chicago Mercantile Exchange; all trades from January 5 to January 16, 1998. Standard deviations and first-order autocorrelations of intertransaction price changes. The price is alternatively measured in logs or levels (for the levels, the units are "ticks").

	Price Variable			
	Log:		Level (Ticks):	
	$\sigma(\Delta\log(P))$ x10,000	$\rho_1(\Delta\log(P))$	$\sigma(\Delta P)$	$\rho_1(\Delta P)$
Feeder Cattle	6.01	-0.18	1.83	-0.18
Live Cattle	4.57	-0.28	1.18	-0.28
Pork Bellies	14.44	-0.10	2.85	-0.10
Australian Dollar	7.12	-0.08	4.60	-0.08
Canadian Dollar	1.51	-0.27	1.06	-0.27
Deutschemark	1.76	-0.19	0.97	-0.19
Japanese Yen	2.04	-0.18	1.56	-0.18
Swiss Franc	1.86	-0.19	1.26	-0.19
13wk Treasury Bill	1.34	-0.04	2.55	-0.04
Eurodollar	0.37	-0.33	0.71	-0.33
Euroyen	0.37	-0.16	0.74	-0.16
One-Month LIBOR	0.38	-0.04	0.72	-0.04
Nasdaq 100	7.83	-0.09	15.61	-0.09
Nikkei 225 Index	11.98	-0.18	3.58	-0.18
S&P 500 Index	1.85	-0.28	1.81	-0.28
S&P Midcap 400 Index	9.92	-0.03	6.36	-0.03

Table 4. Clustering Frequencies $\left(f_\kappa^C\right)$

Sample of contracts traded on the Chicago Mercantile Exchange; all trades from January 5 to January 16, 1998. The clustering frequency is $f_\kappa^C = f_\kappa - (1/\kappa)$ where f_κ is the sample frequency of trades prices that fall on a κ-multiple of the minimum tick. (Since $1/\kappa$ is the expected value under the null hypothesis of uniformly distributed prices, f_κ^C measures "excess" clustering.)

	Tick Multiple κ									
	2	4	5	8	10	20	25	40	50	100
Feeder Cattle	26%	14%	-1%	6%	4%	3%	-1%	2%	0%	0%
Live Cattle	9	6	0	3	1	1	0	0	0	0
Pork Bellies	36	24	1	12	8	6	0	2	2	2
Australian Dollar	5	4	23	3	13	8	5	5	2	0
Canadian Dollar	2	0	1	0	2	1	1	1	1	2
Deutsche Mark	1	1	2	0	1	0	1	0	0	1
Japanese Yen	2	1	4	0	3	1	1	0	1	0
Swiss Franc	1	1	1	0	2	2	1	1	0	0
13wk Treasury Bill	25	8	3	3	8	3	0	2	2	-1
Eurodollar	2	1	1	1	2	-1	-2	-1	-2	-1
Euroyen	9	-7	-1	-6	6	3	-4	0	-2	-1
One-Month LIBOR	9	11	0	1	2	5	3	8	-2	-1
Nasdaq 100	44	40	78	20	82	59	16	29	16	12
Nikkei 225 Index	28	20	12	11	15	9	2	4	1	2
S&P 500 Index	2	1	12	0	7	4	2	2	1	0
S&P Midcap 400 Index	42	22	18	11	23	13	4	5	5	3

Table 5. Incremental Clustering Frequencies $\left(\Delta f_{\kappa}^{C}\right)$

Sample of contracts traded on the Chicago Mercantile Exchange; all trades from January 5 to January 16, 1998. The incremental clustering frequency is constructed to measure clustering after controlling for clustering at the next finer level of resolution: $\Delta f_2^C = f_2 - (1/2) = f_2^C$; $\Delta f_4^C = f_4 - f_2/2$; $\Delta f_5^C = f_5 - (1/5) = f_5^C$; $\Delta f_8^C = f_8 - f_4/2$; $\Delta f_{10}^C = f_{10} - f_5/2$; $\Delta f_{20}^C = f_{20} - f_{10}/2$; $\Delta f_{25}^C = f_{25} - f_5/5$; $\Delta f_{40}^C = f_{40} - f_{20}/2$; $\Delta f_{50}^C = f_{50} - f_{25}/2$; $\Delta f_{100}^C = f_{100} - f_{50}/2$, where f_κ is the sample frequency of trades prices that fall on a κ-multiple of the minimum tick.

	Tick Multiple κ									
	2	4	5	8	10	20	25	40	50	100
Feeder Cattle	26%	1%	-1%	-1%	5%	1%	-1%	0%	1%	0%
Live Cattle	9	2	0	0	2	0	0	-1	0	0
Pork Bellies	36	6	1	-1	8	1	0	0	2	1
Australian Dollar	5	1	23	1	2	1	0	1	0	-1
Canadian Dollar	2	-1	1	0	1	0	1	0	0	1
Deutsche Mark	1	0	2	0	0	0	0	0	0	1
Japanese Yen	2	0	4	-1	1	0	0	-1	0	0
Swiss Franc	1	1	1	0	1	1	1	0	-1	0
13wk Treasury Bill	25	-4	3	-1	6	-1	0	0	2	-2
Eurodollar	2	-1	1	1	1	-2	-2	0	-1	0
Euroyen	9	-11	-1	-3	6	0	-4	-2	0	0
One-Month LIBOR	9	6	0	-5	2	4	3	5	-4	0
Nasdaq 100	44	18	78	0	43	18	0	0	9	3
Nikkei 225 Index	28	6	12	1	9	2	-1	0	1	1
S&P 500 Index	2	0	12	0	1	0	-1	0	0	0
S&P Midcap 400 Index	42	1	18	0	15	1	0	-1	3	0

Table 6. Moment Estimates of the Roll Model

Sample of contracts traded on the Chicago Mercantile Exchange; all trades from January 5 to January 16, 1998. Standard deviations and first-order autocorrelations of intertransaction price changes. The price is alternatively measured in logs or levels (for the levels, the units are "ticks"). σ_u is the standard deviation of the random-walk ("efficient price") component in the model; c is the half-spread.

	Price Variable			
	Log:		Level (Ticks):	
	$\sigma_u \times 10{,}000$	$c \times 10{,}000$	σ_u	c
Feeder Cattle	4.78	2.58	1.46	0.79
Live Cattle	3.02	2.43	0.78	0.63
Pork Bellies	12.97	4.49	2.56	0.88
Australian Dollar	6.54	1.99	4.22	1.29
Canadian Dollar	1.03	0.78	0.72	0.54
Deutschemark	1.40	0.76	0.77	0.42
Japanese Yen	1.64	0.86	1.26	0.65
Swiss Franc	1.47	0.80	1.00	0.54
13wk Treasury Bill	1.28	0.28	2.43	0.53
Eurodollar	0.22	0.22	0.41	0.41
Euroyen	0.31	0.15	0.62	0.29
One-Month LIBOR	0.36	0.08	0.69	0.14
Nasdaq 100	7.13	2.29	14.22	4.56
Nikkei 225 Index	9.65	5.03	2.87	1.51
S&P 500 Index	1.23	0.98	1.20	0.96
S&P Midcap 400 Index	9.63	1.70	6.15	1.16

Table 7. Gibbs Sampler Estimates of the Roll Model.

Gibbs-sampler estimates of price dynamics for the indicated CME contracts, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. σ_u is the implicit efficient price (random-walk) variance; c is the (log) half-spread; $\hat{\pi}^C$ is the probability that a given trade is correctly classified (buy vs. sell). "Mean", "Mode" and "Std.Dev." refer to the posterior distribution; "SEM" is the standard error of the mean, corrected for autocorrelation in the draws.

		Mean	Mode	SEM	Std.Dev.
Pork Belly		Mean	Mode	SEM	Std.Dev.
	$\sigma_u \times 10{,}000$	13.484	13.456	0.005	0.310
	$c \times 10{,}000$	1.092	1.142	0.007	0.426
$\hat{\pi}^C = 55\%$					
Deutschemark		Mean	Mode	SEM	Std.Dev.
	$\sigma_u \times 10{,}000$	1.080	1.080	0.001	0.031
	$c \times 10{,}000$	0.881	0.885	0.001	0.029
$\hat{\pi}^C = 86\%$					
S&P		Mean	Mode	SEM	Std.Dev.
	$\sigma_u \times 10{,}000$	1.629	1.633	0.001	0.045
	$c \times 10{,}000$	0.849	0.842	0.001	0.048
$\hat{\pi}^C = 77\%$					

Table 8. Estimation Results for the Discrete, Clustered Price Model

Notes: Gibbs-sampler estimates of price dynamics for the indicated CME contracts, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. σ_u is the implicit efficient price (random-walk) variance; κ is the (preset) clustering multiple; k is the probability that a transaction price is clustered; $\hat{\pi}^C$ is the probability that a given trade is correctly classified (buy vs. sell). "Mean" , "Mode" and "Std.Dev." refer to the posterior distribution; "SEM" is the standard error of the mean, corrected for autocorrelation in the draws.

		Mean	Mode	SEM	Std.Dev.
Pork Belly					
	$\sigma_u \times 10^4$	11.595	11.570	0.008	0.327
	k	0.739	0.743	0.001	0.021
$\hat{\pi}^C = 71\%$; κ=2					
Deutschemark		Mean	Mode	SEM	Std.Dev.
	$\sigma_u \times 10^4$	1.017	1.014	0.001	0.037
	k	0.003	0.000	0.000	0.004
$\hat{\pi}^C = 81\%$; κ=2					
S&P		Mean	Mode	SEM	Std.Dev.
	$\sigma_u \times 10^4$	1.545	1.532	0.001	0.049
	k	0.043	0.040	0.000	0.011
$\hat{\pi}^C = 70\%$; κ=5					

Table 9. Estimation Results for the Asymmetric Information Model

Notes: Gibbs-sampler estimates of price dynamics for the indicated CME contracts, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. σ_u is the residual (non trade) variance; κ is the (preset) clustering multiple; k is the probability that a transaction price is clustered; the λ's are trade impact coefficients; $\hat{\pi}^C$ is the probability that a given trade is correctly classified (buy vs. sell). "Mean", "Mode" and "Std.Dev." refer to the posterior distribution; "SEM" is the standard error of the mean, corrected for autocorrelation in the draws.

		Mean	Mode	SEM	Std.Dev.
Pork Belly		Mean	Mode	SEM	Std.Dev.
$\hat{\pi}^C = 81\%;\ \kappa = 2$	$\sigma_u \times 10^4$	9.8113	9.7544	0.0216	0.4751
	k	0.7396	0.7417	0.0006	0.0209
	$\lambda_{0,Const} \times 10^4$	0.0490	0.0141	0.0010	0.0427
	$\lambda_{0,\sqrt{}} \times 10^4$	0.1875	0.1926	0.0011	0.0329
	$\lambda_{1,Const} \times 10^4$	0.0558	0.0169	0.0014	0.0512
	$\lambda_{1,\sqrt{}} \times 10^4$	0.1208	0.1190	0.0009	0.0322
Deutschemark		Mean	Mode	SEM	Std.Dev.
$\hat{\pi}^C = 88\%;\ \kappa = 2$	$\sigma_u \times 10^4$	0.7921	0.7831	0.0018	0.0401
	k	0.0038	0.0007	0.0002	0.0050
	$\lambda_{0,Const} \times 10^4$	0.0051	0.0018	0.0001	0.0038
	$\lambda_{0,\sqrt{}} \times 10^4$	0.0048	0.0051	0.0000	0.0010
	$\lambda_{1,Const} \times 10^4$	0.0341	0.0347	0.0004	0.0088
	$\lambda_{1,\sqrt{}} \times 10^4$	0.0050	0.0048	0.0001	0.0017
S&P		Mean	Mode	SEM	Std.Dev.
$\hat{\pi}^C = 72\%;\ \kappa = 5$	$\sigma_u \times 10^4$	1.5271	1.5287	0.0014	0.0570
	k	0.0417	0.0382	0.0004	0.0112
	$\lambda_{0,Const} \times 10^4$	0.0107	0.0032	0.0002	0.0081
	$\lambda_{0,\sqrt{}} \times 10^4$	0.0019	0.0005	0.0000	0.0016
	$\lambda_{1,Const} \times 10^4$	0.0058	0.0017	0.0001	0.0053
	$\lambda_{1,\sqrt{}} \times 10^4$	0.0022	0.0019	0.0000	0.0015

Table 10. Derived Statistics for Asymmetric Information Model

Notes: Gibbs-sampler estimates of price dynamics for the indicated CME contracts, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. In the model *m* is the implicit (log) efficient price; the table summarizes the sources of its variance. "Mean", "Mode" and "Std.Dev." refer to the posterior distribution; "SEM" is the standard error of the mean, corrected for autocorrelation in the draws.

Pork Belly	Contribution to $Var(\Delta m_t)\times 10^6$	Mean	Mode	SEM	Std.Dev.
	Non-trade	0.9649	0.9441	0.0043	0.0940
	Trade	0.2604	0.2655	0.0019	0.0515
	Total	1.2252	1.2011	0.0031	0.0821
Deutschemark	Contribution to $Var(\Delta m_t)\times 10^6$	Mean	Mode	SEM	Std.Dev.
	Non-trade	0.0063	0.0061	0.0000	0.0006
	Trade	0.0039	0.0039	0.0000	0.0006
	Total	0.0101	0.0100	0.0000	0.0007
S&P	Contribution to $Var(\Delta m_t)\times 10^6$	Mean	Mode	SEM	Std.Dev.
	Non-trade	0.0234	0.0233	0.0000	0.0017
	Trade	0.0007	0.0005	0.0000	0.0004
	Total	0.0241	0.0239	0.0000	0.0017

Figure 1. Gibbs Sampler Results for the Basic Roll Model, Pork Belly Contract.

Gibbs-sampler estimates of price dynamics for the pork belly contract, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. The estimated model has two parameters, the log random-walk standard deviation, $\sigma_u \times 10{,}000$ ("SD_u") and the log half-spread $c \times 10{,}000$. The left-most graph plots the actual draws (every tenth draw); the center graph is the autocorrelogram; the right-most graph is the distribution histogram.

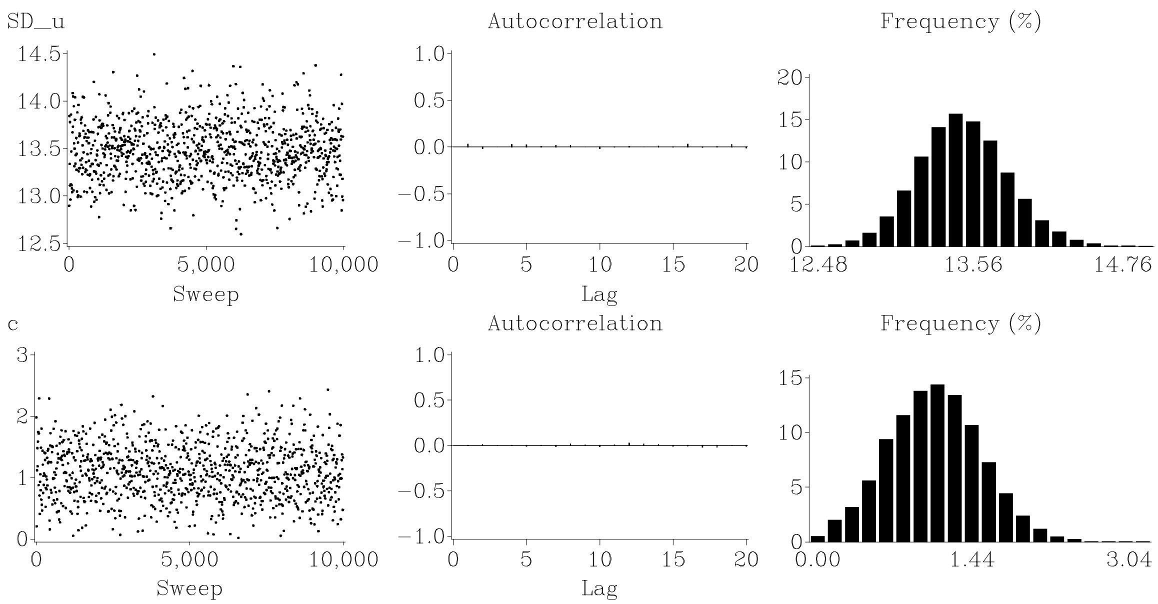

Figure 2. Gibbs Sampler Results for the Discrete Price Model, Pork Belly Contract

Gibbs-sampler estimates of price dynamics for the pork belly contract, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. The estimated model has two parameters, the log random-walk standard deviation, $\sigma_u \times 10{,}000$ ("SD_u") and the mean of the half-spread c. The left-most graph plots the actual draws (every tenth draw); the center graph is the autocorrelogram; the right-most graph is the distribution histogram.

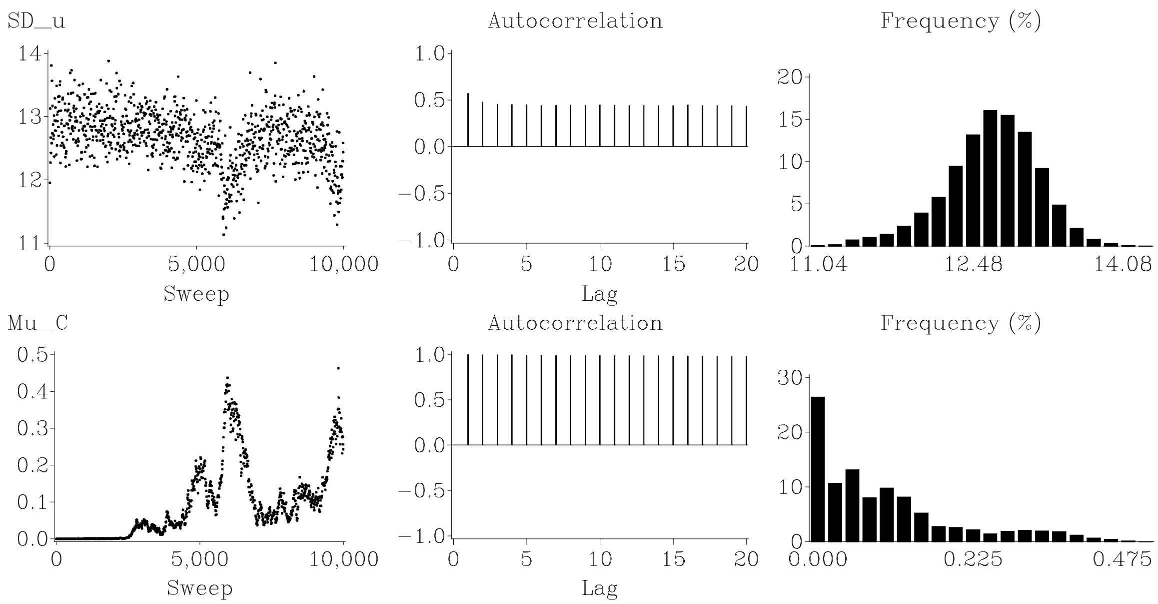

Figure 3. Implied Trade Price Impacts

Notes: Figures are based Gibbs-sampler estimates of price dynamics for the indicated contracts, first 1,000 observations in the two-week sample January 5, 1998 through January 16, 1998. Each figure depicts the contemporaneous and cumulative (through the first lag) impact of a purchase on the implicit log efficient price x 100 (i.e., the vertical units are approximately percentage changes).