Econometrica, Vol. 72, No. 2 (March, 2004), 433–465

INFORMATION IN SECURITIES MARKETS: KYLE MEETS GLOSTEN AND MILGROM

BY KERRY BACK AND SHMUEL BARUCH

This paper analyzes models of securities markets with a single strategic informed trader and competitive market makers. In one version, uninformed trades arrive as a Brownian motion and market makers see only the order imbalance, as in Kyle (1985). In the other version, uninformed trades arrive as a Poisson process and market makers see individual trades. This is similar to the Glosten–Milgrom (1985) model, except that we allow the informed trader to optimize his times of trading. We show there is an equilibrium in the Glosten–Milgrom-type model in which the informed trader plays a mixed strategy (a point process with stochastic intensity). In this equilibrium, informed and uninformed trades arrive probabilistically, as Glosten and Milgrom assume. We study a sequence of such markets in which uninformed trades become smaller and arrive more frequently, approximating a Brownian motion. We show that the equilibria of the Glosten–Milgrom model converge to the equilibrium of the Kyle model.

KEYWORDS: Market microstructure, Kyle model, Glosten–Milgrom model, asymmetric information, insider trading, bid-ask spread, liquidity, depth, bluffing.

THE MOTIVATION for this paper is to understand the relationship between the two canonical models of market microstructure, due to Kyle (1985) and Glosten and Milgrom (1985). The Kyle model is a model of a batch-auction market, in which market makers see only the order imbalance at each auction date. Market makers compete to fill the order imbalance, and matching orders are crossed at the market-clearing price. Because orders are batched, there are no real bid and ask prices. In the Glosten–Milgrom model, orders arrive and are executed by market makers individually. In this model, there are bid and ask quotes, which are determined by the probability that a particular order is informed. Glosten and Milgrom assume the arrival rates of informed and uninformed traders are determined exogenously. Informed traders trade when chosen by this exogenous mechanism as if they have no future opportunities to trade. In other words, when trade is profitable, they trade as much as possible whenever possible. On the other hand, Kyle determines the optimal trading behavior for the single informed agent in his model and shows that in equilibrium the agent will trade on his information only gradually, rather than exploiting it to the maximum extent possible as soon as possible.

The first contribution of this paper is to solve a version of the Glosten–Milgrom model with a single informed trader, in which the informed trader chooses his trading times optimally. We wish to give the trader a rich menu of times from which to choose, and we wish to avoid the likelihood of informed and uninformed trades arriving simultaneously, because the spirit of the model is that market makers execute orders individually. So, we work in continuous time. We assume uninformed trades arrive as Poisson processes. We show that the informed trader plays a mixed strategy, randomizing at each instant between trading and waiting (specifically, his equilibrium strategy is a point process with stochastic intensity). This means that informed and uninformed trades do arrive probabilistically, as Glosten and Milgrom assume. The bid-ask spread, which depends on the relative arrival rates of informed and uninformed trades, is endogenous in our model.

The second contribution of the paper is to show that this version of the Glosten–Milgrom model is approximately the same as the continuous-time version of the Kyle model, when the trade size is small and uninformed trades arrive frequently. The distinction between batching orders, as in Kyle's model, or executing individual orders, as in Glosten and Milgrom's model, turns out to be unimportant when orders are small and arrive frequently. We show that the trading strategies and profits of the informed trader and the losses of uninformed traders are approximately the same in the two models. We also show that the bid-ask spread in the Glosten–Milgrom model is approximately twice the order size multiplied by "Kyle's lambda."

Kyle shows that the equilibrium in his continuous-time model is the limit of discrete-time equilibria of a batch-auction market when the time intervals between trades and the variance of uninformed trades at each date become small. Therefore, our convergence result shows that the discrete-time Kyle model is approximately the same as our version of the Glosten–Milgrom model, when the time intervals and the variance of uninformed trades are small in the former, and the frequency of uninformed trades is large and the order size small in the latter.

The convergence of the informed trading strategy in the Glosten–Milgrom model to that in the Kyle model means that the frequency of informed trades in the Glosten–Milgrom model does not increase at the same rate as the frequency of uninformed trades, when the latter approaches infinity. In the limit, the informed trading strategy is absolutely continuous ("of order dt") whereas the uninformed trades are a Brownian motion ("of order $\sqrt{dt}$"), even though the order size is the same for both types of traders in the Glosten–Milgrom model. In Kyle's derivation of the continuous-time equilibrium as a limit of discrete-time equilibria, this difference in uninformed volume versus informed volume is the result of the informed order being significantly smaller than the variance of uninformed trades at each date. Our results provide an alternative interpretation for the difference in volume: even though the order size for both types of traders is the same, the informed trader simply chooses to trade at a frequency that is lower than the frequency at which uninformed trades arrive.

We focus on mixed-strategy equilibria in the Glosten–Milgrom model because pure-strategy equilibria cannot exist. We assume the asset value $\tilde{v}$ has a Bernoulli distribution with values normalized to zero and one. A pure strategy for the informed trader who knows $\tilde{v} = 1$ would be a sequence of stopping times $\tau_1, \tau_2, \ldots$, measurable with respect to public information, at which he buys or sells the asset. If in equilibrium the informed trader were to follow such a strategy and there were no order at time τ_1, then market makers would assume that $\tilde{v} = 0$ and would quote bids and asks at 0 thereafter. Consequently, when τ_1 is reached, the optimal strategy would be to abstain from the market at that instant and then buy an infinite amount beginning immediately afterward. We conclude that there can be no equilibrium in pure strategies.

In a mixed strategy equilibrium, an agent is of course indifferent among the various choices over which he randomizes. In our model these choices are to trade (either to buy or to sell or he may randomize over both) or to wait to trade. The willingness of the informed trader to wait to trade in this version of the Glosten–Milgrom model is analogous to his trading gradually on his information in Kyle's model.

We show that in certain circumstances the informed trader in the Glosten–Milgrom model will randomize over all the alternatives available to him, including trading in the "wrong" direction. We call this phenomenon "bluffing." Black (1990) discusses bluffing by uninformed traders who trade as if they had news, using market orders, and then reverse their trades using limit orders. We do not allow the informed trader to use limit orders; nevertheless, he sometimes trades as if he had the opposite information. Huddart, Hughes, and Levine (2001) also show that informed traders may bluff, but in their model, traders are required to disclose their trades ex-post, and it is this disclosure that makes it profitable to bluff. We show that bluffing can occur even when trades are ex-post anonymous. We should emphasize that the trader does not profit from bluffing, so it is perhaps incorrect to think of it as "manipulating" the market. The trader is simply indifferent in equilibrium between trading in the normal direction, waiting to trade, and bluffing. Each of the latter alternatives increases the potential for future profits at the expense of current profits.

The Kyle version of our model is analyzed in Section 1 and the Glosten–Milgrom version in Section 2. Section 3 establishes the convergence result, and Section 4 documents that there is sometimes bluffing in the Glosten–Milgrom model. Section 5 concludes the paper.

1. KYLE MODEL

We consider a continuous-time market for a risky asset and one risk-free asset with interest rate set to zero. A public release of information takes place at a random time τ, distributed as an exponential random variable with parameter r. After the public announcement has been made, the value of the risky asset, denoted by $\tilde{v}$, will be either zero or one, and all positions are liquidated at that price. There is a single informed trader who knows $\tilde{v}$ at date 0. There are also uninformed (presumably liquidity motivated) trades that arrive as a Brownian motion with volatility σ^2. We will call these "noise trades." All trades are anonymous. Competitive risk neutral market makers absorb the net order flow, the competition ensuring that the transaction price is always the conditional expectation of $\tilde{v}$, given the information in orders. The informed trader recognizes that his trades affect prices through the Bayesian updating of market makers. Our goal is to determine the optimal trading behavior of the informed trader and the equilibrium price adjustments, as a function of order flow, of market makers.

This model differs from the continuous-time model of Kyle (1985) in only two ways: the announcement date is random, and the asset value is zero or one, rather than normally distributed. The assumption of a random announcement date may be more or less reasonable than assuming a known announcement date, depending on the context. Our motive for the assumption is tractability: it means that we can eliminate time as a state variable. Our motive for the distribution assumption on the asset value is that it simplifies the Glosten–Milgrom-type model. It is actually a bit of a hindrance in the Kyle-type model.[1]

Without loss of generality, we take the initial position of the informed trader in the risky asset to be zero. We denote by X_t the number of shares held at time t. We let Z_t denote the number of shares held by noise traders at date t, taking $Z_0 = 0$. We require that X be adapted to the filtration generated by Z and $\tilde{v}$. This does not literally require that Z be observed by the informed trader, because, given strictly monotone pricing, Z can be inferred by the informed trader from the price process.

Setting $Y = X + Z$, the net order received by market makers at time t can be viewed as $dY_t \equiv dX_t + dZ_t$. Risk neutrality and competition between the market makers implies that the price at time $t < \tau$ is

$$p_t = E[\tilde{v} | (Y_s)_{s \leq t}]. \tag{1.1}$$

The initial price p_0 is the unconditional expectation of $\tilde{v}$, and we assume $0 < p_0 < 1$.

We assume that[2]

$$X_t = \int_0^t \theta_s \, ds \tag{1.2}$$

for some stochastic process θ (depending on $\tilde{v}$). We will show that there is an equilibrium in which the rate of trade of the informed trader at time t is only a function of p_t and $\tilde{v}$. We will denote the rate of trade of the high-type informed trader by $\theta_H(p_t)$ and the rate of trade of the low type by $-\theta_L(p_t)$. We will show that $\theta_H > 0$ and $\theta_L > 0$, so θ_H denotes the rate at which the high type buys the security and θ_L denotes the rate at which the low type sells. Whether the trader has good or bad news, his order rate in this circumstance is

$$\theta_t = \tilde{v}\theta_H(p_t) - (1 - \tilde{v})\theta_L(p_t). \tag{1.3}$$

[1]Back (1992) solves the Kyle model for more general distributions than the normal, but he still requires the distribution to be continuous. We do not know of any literature on the continuous-time Kyle model with discrete distributions.

[2]Back (1992) shows that optimal strategies in Kyle models have this property.

In this case, the high-type informed trader's expected profit is

$$(1.4) \qquad E\int_0^\tau (1-p_t)\theta_H(p_t)\,dt,$$

and the low type's expected profit is

$$(1.5) \qquad E\int_0^\tau p_t\theta_L(p_t)\,dt.$$

When the informed trader's strategy is as described in the previous paragraph, the competitive pricing assumption (1.1) implies a specific form for the price process. Given (1.3), the expected rate of informed trade at time t, given the market makers' information, is

$$p_t\theta_H(p_t) - (1-p_t)\theta_L(p_t).$$

Hence, the surprise or innovation in the order flow at time t is

$$dY_t - \phi(p_t)\,dt,$$

where

$$(1.6) \qquad \phi(p) \equiv p\theta_H(p) - (1-p)\theta_L(p).$$

In this model, we have the standard result that the revision in beliefs is proportional to the surprise in the variable being observed, which means that

$$(1.7) \qquad dp_t = \lambda(p_t)[dY_t - \phi(p_t)\,dt],$$

for some function λ. Fairly standard filtering theory (we derive this in the Appendix, in the proof of Theorem 1) shows in fact that, for $0 < p < 1$,

$$(1.8) \qquad \lambda(p) = \frac{p(1-p)[\theta_H(p) + \theta_L(p)]}{\sigma^2}.$$

We need to augment (1.7) by defining the boundary behavior at 0 and 1. New information will not change beliefs that put probability zero or one on the asset value being high, so we specify that 0 and 1 are absorbing states for p.

We assume that the informed trader believes prices evolve in accordance with (1.7), for some functions λ and ϕ that he takes as given. In general—i.e., without imposing (1.3)—the expected profit of the informed trader is

$$E\left[\int_0^\tau (\tilde{v} - p_t)\theta_t\,dt \,\middle|\, \tilde{v}\right].$$

We need to rule out strategies that first incur infinite losses and then reap infinite profits, because in that case the overall profit is undefined.[3] In fact, we need to ensure that expected profits are well defined, so we require expected total losses to be finite. We define a strategy $dX_t = \theta_t\,dt$ to be admissible for the high type if

$$E\int_0^\tau (1-p_t)\theta_t^-\,dt < \infty, \tag{1.9a}$$

where as usual $\theta_t^- = \max(0, -\theta_t)$ and we define it to be admissible for the low type if

$$E\int_0^\tau p_t\theta_t^+\,dt < \infty, \tag{1.9b}$$

where $\theta_t^+ = \max(0, \theta_t)$. In these constraints, the price process p is to be understood as generated by (1.7) for given functions λ and ϕ; hence, it depends on θ.

We now define an equilibrium to be a collection $\{\lambda, \phi, \theta_H, \theta_L\}$ such that λ and ϕ are locally Lipschitz[4] and:

(a) given the informed trading strategy (1.3), the solution to the price dynamics (1.7), with absorbing boundaries at 0 and 1, satisfies the competitive pricing condition (1.1);

(b) given the price dynamics (1.7), with absorbing boundaries at 0 and 1, the informed trader's strategy (1.3) maximizes his expected profit over all absolutely continuous strategies $dX_t = \theta_t\,dt$ satisfying the admissibility constraint (1.9).

All proofs are provided in the Appendix. A key result in the proof of the following is that, as in Kyle (1985), beliefs of market makers converge over time to the truth.

THEOREM 1: *Let $N(\cdot)$ denote the standard normal distribution function. For each $0 < p < 1$, define $\lambda^*(p)$ by*

$$N\left(-\sqrt{\log\left(\frac{r}{\pi\sigma^2}\right) - 2\log(\lambda^*(p))}\right) = \min\{p, 1-p\}. \tag{1.10}$$

[3]This is a meaningful restriction, because it will be possible for the high type to push the price to zero and then buy an infinite amount at a zero price; however doing so will first generate an infinite loss. There will be no strategies that generate an infinite gain without first generating an infinite loss.

[4]This guarantees the existence of a unique strong solution to (1.7), since the absorbing boundaries rule out the possibility of a finite explosion time. See Protter (1990, p. 199).

Set $\lambda^*(0) = \lambda^*(1) = 0$. *Set* $\phi^* = 0$. *Define the buying rate of the high-type informed trader by*

$$\theta_H^*(p) = \frac{\sigma^2 \lambda^*(p)}{p} \tag{1.11}$$

and the selling rate of the low-type informed trader by

$$\theta_L^*(p) = \frac{\sigma^2 \lambda^*(p)}{1-p}. \tag{1.12}$$

Then $\{\lambda^*, \phi^*, \theta_H^*, \theta_L^*\}$ *is an equilibrium. At any date* t *prior to the announcement date* τ, *given the price* $p_t = p$ *at date* t, *the maximum expected profit achievable by the high-type informed trader during the period* $[t, \tau]$ *is*

$$\int_p^1 \frac{1-a}{\lambda^*(a)}\, da, \tag{1.13}$$

and the maximum expected profit achievable by the low-type informed trader during the period $[t, \tau]$ *is*

$$\int_0^p \frac{a}{\lambda^*(a)}\, da. \tag{1.14}$$

Moreover, λ^* *is a continuous concave function on* $[0, 1]$, *symmetric about* $p = 1/2$, *and attaining a maximum at* $p = 1/2$ *equal to* $\sqrt{r/(\pi\sigma^2)}$.

Concavity of λ is an interesting property, because it explains the patience of the informed trader. In the model of Kyle (1985), the informed trader trades gradually as here, and λ is constant. The main difference between the models is that in this model the trader faces the risk of losing his informational advantage with probability $r\,dt$ in each instant dt. Nevertheless, he is willing to trade gradually instead of capitalizing immediately on his informational advantage. The reason is that, if he abstains from trade, $\lambda(p_t)$ will be a supermartingale, decreasing in expectation over time. Thus, the market becomes deeper on average when he abstains from trade, and this provides sufficient benefit to offset the risk of waiting to trade. The supermartingale property follows from concavity of $\lambda(\cdot)$, Jensen's inequality, and the fact that the price process p is a martingale relative to his information set when the informed trader abstains from trading. The higher is r, the higher is the risk of waiting to trade, hence the greater must be the expected decrease in λ when the informed trader abstains from trading. The theorem is consistent with this, because the degree of concavity of λ is determined by its peak $\sqrt{r/(\pi\sigma^2)}$.

The formulas (1.11) and (1.12) follow from the filtering equations (1.6) and (1.8) when we set $\phi = 0$. Satisfying the filtering equations implies that the competitive pricing condition (1.1) holds. The formulas (1.13) and (1.14) for the value functions of the informed trader can be interpreted as follows (we give a different, rigorous treatment in the Appendix). Consider the high type and consider the strategy of purchasing an arbitrarily large quantity in an arbitarily short period of time. Given that the trader is indifferent between trading and waiting, this should be an optimal strategy. Ignoring the noise trades that arrive in this short period of time, the informed trader will move up the market supply curve $dp = \lambda(p)\,dx$, generating profit

$$(1 - p)\,dx = \frac{1 - p}{\lambda(p)}\,dp.$$

The formula (1.13) is simply the cumulative profit, starting at the given price p and purchasing the asset until $p = 1$, at which point there is no further profit to be earned.

The key to deriving the equilibrium is the analysis of the informed trader's optimization problem. In the proof, we do not assume the Bellman equation (we provide what is called a "verification theorem"). However, we did assume it when originally deriving the formulas in the theorem, and we will sketch our analysis here, because it provides some insight into the theorem. Denote the value function for the high-type informed trader by V and the value function for the low type by J. The stationarity of the problem facing the informed trader implies that these are strictly functions of the price p. Consider the high type. If he trades at rate θ then the expected change in p given his information is $\lambda\theta\,dt - \lambda\phi\,dt$, the volatility of p is $\sigma^2\lambda^2\,dt$, and the probability that V drops to zero is $r\,dt$. Therefore, the expected change in the value function is

$$-rV\,dt + \lambda(\theta - \phi)V'\,dt + \frac{1}{2}\sigma^2\lambda^2 V''\,dt.$$

The Bellman equation states that the sum of this and the expected instantaneous profit from trade, namely $(1 - p)\theta\,dt$, should be zero at the optimal θ. Because this sum is linear in θ, the maximum can be zero only if the coefficient of θ is zero and the remaining terms sum to zero. This means

$$V' = \frac{p - 1}{\lambda}, \tag{1.15}$$

$$rV = -\lambda\phi V' + \frac{1}{2}\sigma^2\lambda^2 V''. \tag{1.16}$$

Similarly, for the low type we should have

$$J' = \frac{p}{\lambda}, \tag{1.17}$$

$$rJ = -\lambda\phi J' + \frac{1}{2}\sigma^2\lambda^2 J''. \tag{1.18}$$

The boundary conditions $\lambda(0) = \lambda(1) = 0$ are natural to impose, given that 0 and 1 must be absorbing states. Assuming enough differentiability, the conditions (1.15)–(1.18) and these boundary conditions determine ϕ, λ, V, and J uniquely. To put this another way, there is only one ϕ and λ for which the conditions (1.15)–(1.18) admit a solution in V and J, and hence only one ϕ and λ for which the informed trader has an optimal trading strategy (assuming, in addition to enough differentiability, that the Bellman equation is a necessary condition for optimality). The key to analyzing (1.15)–(1.18) is to observe that if we differentiate both sides of (1.16) we obtain an equation in ϕ and λ and their derivatives and in V', V'', and V'''. Each of the derivatives of V can be calculated in terms of λ and its derivatives from (1.15), so we can eliminate V', V'', and V''' and obtain a differential equation in ϕ and λ, namely

$$-\phi + (1-p)\phi' + \frac{1}{2}\sigma^2(1-p)\lambda'' = \frac{(p-1)r}{\lambda}. \tag{1.19}$$

Similarly, (1.17) and (1.18) imply the following differential equation in ϕ and λ:

$$-\phi - p\phi' - \frac{1}{2}\sigma^2 p\lambda'' = \frac{pr}{\lambda}. \tag{1.20}$$

These two equations have a unique solution, obtained as follows. Simply subtract (1.20) from (1.19) to obtain

$$\phi' + \frac{1}{2}\sigma^2\lambda'' + \frac{r}{\lambda} = 0.$$

Rearranging (1.20) also gives

$$\phi' + \frac{1}{2}\sigma^2\lambda'' + \frac{r}{\lambda} = -\frac{\phi}{p},$$

so we conclude that $\phi = 0$. With $\phi = 0$, (1.19) and (1.20) are each equivalent to

$$\frac{1}{2}\sigma^2\lambda'' = -\frac{r}{\lambda}. \tag{1.21}$$

We solved (1.21) in conjunction with the boundary conditions $\lambda(0) = \lambda(1) = 0$ by separation of variables to obtain the formula (1.10) for λ^*.

The collection $\{\lambda^*, \phi^*, \theta_H^*, \theta_L^*\}$ should actually be the unique equilibrium, under assumptions sufficient to guarantee the requisite differentiability and sufficient to guarantee that the Bellman equation is necessary for optimality. A result of this sort is given in Back (1992).

2. GLOSTEN–MILGROM MODEL

We consider a model similar to Glosten and Milgrom (1985), except that trades take place at random dates. As in Easley, Kiefer, O'Hara, and Paperman (1996), we assume uninformed buy and sell orders arrive as Poisson processes with constant, exogenously given, arrival intensities β. However, we will endogenize the arrival intensity of informed trades.

We denote the order size by δ. We denote the total number of buy orders by noise traders through time t by z_t^+ and the total number of sell orders by noise traders through time t by z_t^-, and we set $z_t = z_t^+ - z_t^-$. The net number of shares bought by noise traders is then $z_t\delta$. Similarly, we denote the number of informed buys by x_t^+, the number of informed sells by x_t^- and the net informed orders by $x_t = x_t^+ - x_t^-$. Finally, we denote the number of net noise and informed orders by $y_t = z_t + x_t$. The process y reveals the complete history of anonymous trades. The σ-field generated by $\{y_s \mid s \le t\}$ is denoted by $\mathcal{F}_t^y$. As usual, we denote the left limit of y at time t by y_{t-} and set $\Delta y_t = y_t - y_{t-}$. If there is a buy order at date t then $\Delta y_t = 1$, and if there is a sell order then $\Delta y_t = -1$.

Competition among the market makers implies that any transaction takes place at price $p_t \equiv E[\tilde{v} | \mathcal{F}_t^y]$. The posted ask and bid prices at time $t < \tau$ are

$$\text{ask}_t = E[\tilde{v} \mid \mathcal{F}_{t-}^y, \Delta y_t = +1],$$

$$\text{bid}_t = E[\tilde{v} \mid \mathcal{F}_{t-}^y, \Delta y_t = -1].$$

Here, $\mathcal{F}_{t-}^y \equiv \bigcup_{s<t} \mathcal{F}_s^y$ denotes the information available to the market makers just before time t. We also denote the left limit of p at time t by p_{t-}. The interpretation of p_{t-} is that it is the probability market makers place on the event $\tilde{v} = 1$, prior to observing whether there is a trade at time t, and hence it is the expected value of $\tilde{v}$ prior to observing the order flow at time t. The ask and bid are of course the posterior probabilities (and expectations), given a buy and sell order respectively.

The informed trader chooses a trading strategy x to maximize

$$E\left[\delta \int_0^\tau [\tilde{v} - \text{ask}_t]\, dx_t^+ + \delta \int_0^\tau [\text{bid}_t - \tilde{v}]\, dx_t^- \middle| \tilde{v}\right]. \tag{2.1}$$

The integral $\int_0^\tau [\tilde{v} - \text{ask}_t]\delta\, dx_t^+$ is the sum until the announcement date τ of the profit per trade $[\tilde{v} - \text{ask}_t]\delta$, summed over the various dates at which the informed trader buys (i.e., the dates at which $dx_t^+ = 1$), and similarly for the integral $\int_0^\tau [\text{bid}_t - \tilde{v}]\delta\, dx_t^-$. It is not entirely obvious that the expectation in (2.1) exists. As in the previous section, to guarantee that expected profits are well defined (in the extended reals), we rule out strategies that incur infinite expected losses. We define a strategy to be admissible if

$$E \int_0^\tau [1 - \text{bid}_t]\, dx_t^- < \infty \tag{2.2a}$$

for the high-type trader and

$$E\int_0^\tau \text{ask}_t\, dx_t^+ < \infty \tag{2.2b}$$

for the low type.

We assume the informed trader buys or sells at most one unit (i.e., δ shares) at any point in time, since to trade multiple units would identify the trades as being informed. We search for an equilibrium in which the high-type informed trader buys the security in the instant dt with probability $\theta_{HB}(p_{t-})\,dt$ and sells with probability $\theta_{HS}(p_{t-})\,dt$ and the low type buys and sells with probabilities $\theta_{LB}(p_{t-})\,dt$ and $\theta_{LS}(p_{t-})\,dt$ respectively. Specifically, we search for θ's such that the stochastic processes

$$x_t^+ - \tilde{v}\int_0^t \theta_{HB}(p_{s-})\,ds - (1-\tilde{v})\int_0^t \theta_{LB}(p_{s-})\,ds \tag{2.3}$$

and

$$x_t^- - \tilde{v}\int_0^t \theta_{HS}(p_{s-})\,ds - (1-\tilde{v})\int_0^t \theta_{LS}(p_{s-})\,ds \tag{2.4}$$

are martingales, relative to the informed trader's information. In this circumstance, the expected profit is

$$E\int_0^\tau [1-\text{ask}_t]\,\delta\theta_{HB}(p_{t-})\,dt - E\int_0^\tau [1-\text{bid}_t]\,\delta\theta_{HS}(p_{t-})\,dt$$

for the high-type informed trader and

$$-E\int_0^\tau \text{ask}_t\;\delta\theta_{LB}(p_{t-})\,dt + E\int_0^\tau \text{bid}_t\;\delta\theta_{LS}(p_{t-})\,dt$$

for the low type.

We will now solve for the bid and ask prices, assuming (2.3) and (2.4) are martingales. The probability of a buy order arriving in an instant dt is $p\theta_{HB}(p)\,dt + (1-p)\theta_{LB}(p)\,dt + \beta\,dt$, where the three terms refer to the three possible sources of an order: the high-type informed trader, the low-type informed trader, and uninformed traders. The ask price is the sum of the value-weighted conditional probabilities of the order coming from each of the three possible sources, namely,

$$\frac{p\theta_{HB}}{p\theta_{HB}+(1-p)\theta_{LB}+\beta}\times 1 + \frac{(1-p)\theta_{LB}}{p\theta_{HB}+(1-p)\theta_{LB}+\beta}\times 0 + \frac{\beta}{p\theta_{HB}+(1-p)\theta_{LB}+\beta}\times p.$$

Thus, the ask price at time t is $a(p_{t-})$, where

$$a(p) \equiv \frac{p\theta_{HB}(p) + p\beta}{p\theta_{HB}(p) + (1-p)\theta_{LB}(p) + \beta}. \tag{2.5}$$

Similarly, the bid price must be $b(p_{t-})$, where

$$b(p) \equiv \frac{p\theta_{HS}(p) + p\beta}{p\theta_{HS}(p) + (1-p)\theta_{LS}(p) + \beta}. \tag{2.6}$$

Observe that to have $a > p$ we must have $\theta_{HB} > \theta_{LB}$ and in particular $\theta_{HB} > 0$. Similarly, $b < p$ implies $\theta_{LS} > \theta_{HS} \geq 0$.

Considering only the jumps to the ask or the bid, the expected change in p in an instant dt is

$$\begin{aligned}
&(a(p) - p)[p\theta_{HB} + (1-p)\theta_{LB} + \beta]\,dt \\
&\quad + (b(p) - p)[p\theta_{HS} + (1-p)\theta_{LS} + \beta]\,dt \\
&= p(1-p)[\theta_{HB} + \theta_{HS} - \theta_{LB} - \theta_{LS}]\,dt.
\end{aligned}$$

The process p is a martingale, so this expected change must be canceled by the expected change in p between orders. We conclude that between orders $dp_t = f(p_{t-})\,dt$, where

$$f(p) \equiv p(1-p)[\theta_{LB}(p) + \theta_{LS}(p) - \theta_{HB}(p) - \theta_{HS}(p)]. \tag{2.7}$$

Summarizing, the evolution of p must be given by[5]

$$dp_t = f(p_{t-})\,dt + [a(p_{t-}) - p_{t-}]\,dy_t^+ + [b(p_{t-}) - p_{t-}]\,dy_t^-. \tag{2.8}$$

As in the previous section, we take 0 and 1 to be absorbing points for p.

The informed trader takes f, a, and b to be given functions and maximizes his expected profit subject to (2.8) and the boundary conditions that 0 and 1 are absorbing. As in the previous section, we denote the value function for the high-type informed trader by V and the value function for the low type by J. Because τ is exponentially distributed, the value functions depend only on p_{t-} for $t < \tau$ (they jump to zero at τ).

As observed above, to have $a > p$ and $b < p$, we must have $\theta_{HB} > 0$ and $\theta_{LS} > 0$, which means that it must be optimal for the high type to buy and the low type to sell at each time t. Because p jumps to $a(p)$ when there is a buy

[5]The existence of a unique strong solution to this stochastic differential equation is guaranteed by the differentiability assumption in Theorem 2, which implies that f, a, and b are continuously differentiable, hence locally Lipschitz. See Protter (1990, p. 199).

order and falls to $b(p)$ when there is a sell order, this optimality implies

$$V(p) = [1 - a(p)]\delta + V(a(p)), \tag{2.9}$$

$$J(p) = b(p)\delta + J(b(p)). \tag{2.10}$$

Also, it can never be better than optimal for the high type to sell or the low type to buy, and if the high type sells with positive probability or the low type buys with positive probability, then it must be optimal to do so. This is equivalent to

$$V(p) \geq [b(p) - 1]\delta + V(b(p)), \quad \text{with equality when } \theta_{HS} > 0, \tag{2.11}$$

$$J(p) \geq -a(p)\delta + J(a(p)), \quad \text{with equality when } \theta_{LB} > 0. \tag{2.12}$$

It must also be optimal to refrain from trading at each time. During an instant dt the announcement occurs with probability $r\,dt$ and if the announcement occurs, then the value function becomes 0. An uninformed buy order will arrive with probability $\beta\,dt$, in which case p will jump to $a(p)$ and the value function for the high type will jump to $V(a(p))$. Similarly, with probability $\beta\,dt$ the value function will jump to $V(b(p))$. Finally in the absence of an announcement or an order, p will change by $f(p)\,dt$ and the value function will change by $V'(p)f(p)\,dt$. For it to be optimal for the high type to refrain from trading, all of these expected changes in V must cancel. Applying this logic to the low-type informed trader also, we have

$$rV(p) = V'(p)f(p) + \beta[V(a(p)) - V(p)] + \beta[V(b(p)) - V(p)], \tag{2.13}$$

$$rJ(p) = J'(p)f(p) + \beta[J(a(p)) - J(p)] + \beta[J(b(p)) - J(p)]. \tag{2.14}$$

The natural boundary conditions (which hold also in our version of the Kyle model) are $J(0) = V(1) = 0$ and $J(1) = V(0) = \infty$.

The following establishes the sufficiency of these conditions for equilibrium. The key step in the proof is the replication of the result of Glosten and Milgrom (1985) that beliefs of market makers converge over time to the truth.

THEOREM 2: *Let a, b, f, V, J, θ_{HB}, θ_{HS}, θ_{LB}, and θ_{LS} satisfy* (2.5)–(2.14), *with $\theta_{HB} > \theta_{LB}$ and $\theta_{LS} > \theta_{HS}$. Assume the trading strategies are admissible. Assume V is a nonincreasing and J is a nondecreasing function of p, and V and J satisfy the boundary conditions*

$$\lim_{p\to 0} V(p) = \lim_{p\to 1} J(p) = \infty \quad \textit{and} \quad \lim_{p\to 1} V(p) = \lim_{p\to 0} J(p) = 0.$$

Assume the functions θ_{HS}, θ_{HB}, θ_{LB}, and θ_{LS} are continuously differentiable on $(0, 1)$. Then the informed trading strategy is optimal and, for all t, $p_t = E[\tilde{v}|\mathcal{F}_t^y]$.

We can solve conditions (2.5)–(2.14) numerically. We consider a discrete grid on $(0, 1)$ and use a variation of value iteration to compute the value function at each p in the grid. Our numerical method is explained in Appendix B.

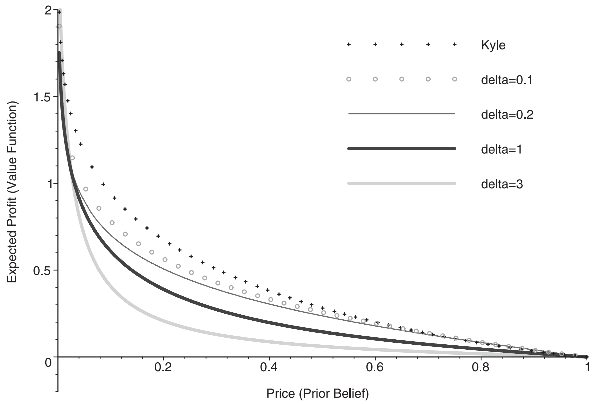

FIGURE 1.—Value functions in the Glosten–Milgrom model. Assuming $r = 1$ and $2\beta\delta^2 = 1$, this shows the value function $V(p)$ of the high-type informed trader in our version of the Glosten–Milgrom model. As δ decreases, the value functions are approaching the value function from our version of the Kyle model (labeled "Kyle").

In Figures 1 and 2 we show the value function V and the bid and ask prices for different values of δ, with in each case $\beta = 1/(2\delta^2)$ and $r = 1$. Figure 1 also shows the value function for our version of the Kyle model with $\sigma = 1$. Figure 1 suggests that the Glosten–Milgrom models converge to the Kyle model as δ becomes smaller. This is verified in the following section.

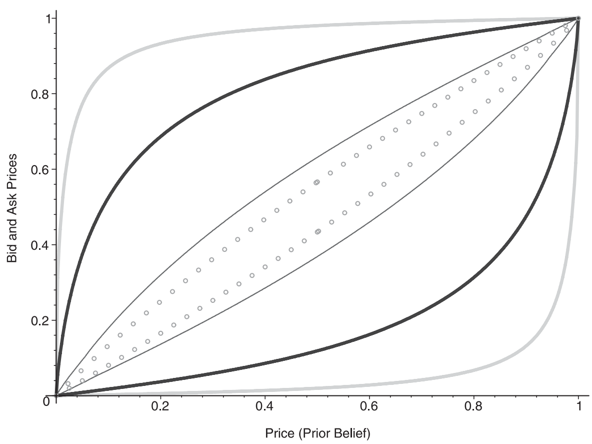

FIGURE 2.—Bid and ask prices in the Glosten–Milgrom model. Assuming $r = 1$ and $2\beta\delta^2 = 1$, this shows the bid and ask prices in our version of the Glosten–Milgrom model as functions of p, for $\delta = 3$, $\delta = 1$, $\delta = .2$, and $\delta = .1$. The upper curves are the ask functions and the lower curves are the bid functions. As δ decreases, the ask functions are decreasing and the bid functions increasing towards the 45° line.

3. CONVERGENCE

We now consider the convergence of the Glosten–Milgrom equilibria to the Kyle equilibrium, when the trade size becomes small ($\delta \to 0$) and the noise trades arrive more frequently ($\beta \to \infty$). We will take $\beta = \sigma^2/(2\delta^2)$. This implies that the expected squared noise trade per unit of time is $2\beta\delta^2 = \sigma^2$, as in the Kyle model. In fact, these assumptions imply that the process of cumulative noise trades in the Glosten–Milgrom model converges weakly to the Brownian motion in the Kyle model (see, e.g., Ethier and Kurtz (1986)).

We assume the existence of an equilibrium in the Glosten–Milgrom model for each δ. We denote the equilibrum value functions by V_δ and J_δ and the equilibrium ask and bid prices by a_δ and b_δ. The equilibrium θ's also depend on δ but our notation will not indicate it.

The following verifies the convergence of the value functions indicated by Figure 1. Part (b) establishes the relationship between the bid-ask spread in the Glosten–Milgrom model and "Kyle's lambda."

THEOREM 3: *Assume the equilibria satisfy* (2.5)–(2.14). *Assume V_δ and J_δ converge in the topology of C^2-convergence on compact subsets of* $(0,1)$ *to functions V and J that are strictly monotone* (V *decreasing and* J *increasing*) *and satisfy the boundary conditions*

$$\lim_{p\to 0} V(p) = \lim_{p\to 1} J(p) = \infty \quad \textit{and} \quad \lim_{p\to 1} V(p) = \lim_{p\to 0} J(p) = 0.$$

Then:

(a) *V and J must be the value functions from our version of the Kyle model defined in* (1.13) *and* (1.14);

(b) *for each* $p \in (0,1)$,

$$\lim_{\delta\to 0} \frac{a_\delta(p) - p}{\delta} = \lim_{\delta\to 0} \frac{p - b_\delta(p)}{\delta} = \lambda^*(p),$$

where λ^* is the market depth parameter (1.10) *from our version of the Kyle model*;

(c) *for each* $p \neq 1/2$, *there exists* $\hat{\delta}(p)$ *such that for all* $\delta \le \hat{\delta}(p)$, *either* $\theta_{LB}(p) > 0$ *or* $\theta_{HS}(p) > 0$.

Part (b) means that for small δ the change in the conditional expectation of the asset value when a buy order arrives is approximately $\delta\lambda^*(p)$, and the change in the conditional expectation upon receipt of a sell order is approximately $-\delta\lambda^*(p)$. This is fully consistent with Kyle's (1985) interpretation of λ as the market depth parameter. Figure 3 illustrates the convergence of $(a_\delta - p)/\delta$ to λ^*.

Part (c) means that for sufficiently small δ, either the low-type informed trader buys the asset with positive probability, or the high-type informed trader

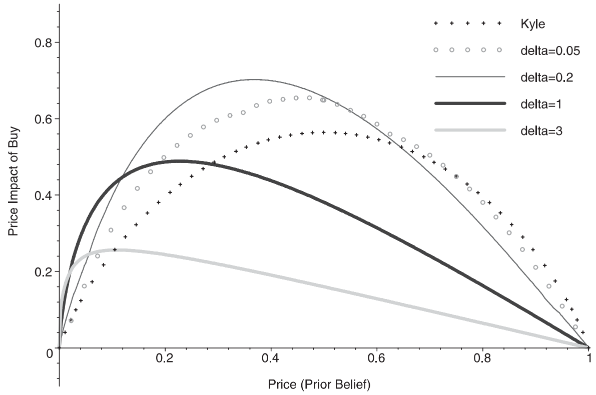

FIGURE 3.—Convergence of the price impact of buy orders. Assuming $r = \sigma = 1$ and $2\beta\delta^2 = 1$, this plots $(a_\delta(p) - p)/\delta$ for different values of δ. The modes of the curves are moving to the right as δ decreases, and the curves are taking on the shape of the plot labeled "Kyle," which is the plot of the market depth parameter $\lambda^*(p)$ from our version of the Kyle model.

sells it with positive probability. We refer to this phenomenon as "bluffing," and we will discuss it further in the next section.

Under the assumption that the intensity of bluffing does not increase at the same rate as the intensity of noise trading, we can establish the convergence of the informed trading strategies and give a more precise result about when each type of trader bluffs.

COROLLARY: *Assume the hypothesis of Theorem* 3. *Assume in addition that* $\theta_{LB}(p)/\beta$ *and* $\theta_{HS}(p)/\beta$ *converge to zero for each* $p \in (0, 1)$ *as* $\delta \to 0$. *Then*:

(a) *for each* $p \in (0, 1)$,

$$\delta[\theta_{HB}(p) - \theta_{HS}(p)] \to \theta^*_H(p),$$

and

$$\delta[\theta_{LS}(p) - \theta_{LB}(p)] \to \theta^*_L(p),$$

where θ^*_H *and* θ^*_L *are the equilibrium order rates in our version of the Kyle model defined in* (1.11) *and* (1.12);

(b) *for each* $p > 1/2$, $\theta_{HS}(p) > 0$ *for all sufficiently small* δ, *and*, *for each* $p < 1/2$, $\theta_{LB}(p) > 0$ *for all sufficiently small* δ.

The additional hypothesis of the corollary is quite weak. We know that $\theta_{LB}(p) < \theta_{HB}(p)$ by virtue of the fact that $a_\delta(p) > p$. Likewise $\theta_{HS}(p) < \theta_{LS}(p)$ by virtue of the fact that $b_\delta(p) < p$. Our numerical results show that in fact θ_{HB}/β and θ_{LS}/β converge to zero. Figure 4 displays the numerical results for θ_{HB}/β.

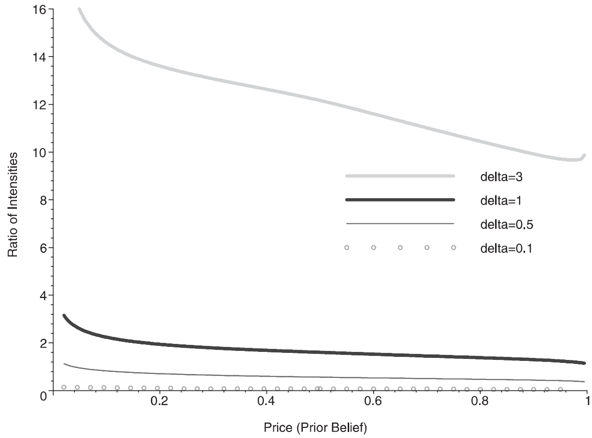

FIGURE 4.—Convergence of relative order intensities. Assuming $r = \sigma = 1$ and $2\beta\delta^2 = 1$, this shows the convergence of θ_{HB}/β to zero as δ decreases.

Figure 5 illustrates part (a) of the Corollary. It shows $\delta(\theta_{HB} - \theta_{HS})$ converging to θ_H^*.[6] The interpretation of $\delta(\theta_{HB} - \theta_{HS})\,dt$ is that it is the expected net buy order of the high type in the Glosten–Milgrom model in an instant dt. The analogous concept in the Kyle model is $\theta_H^*\,dt$.

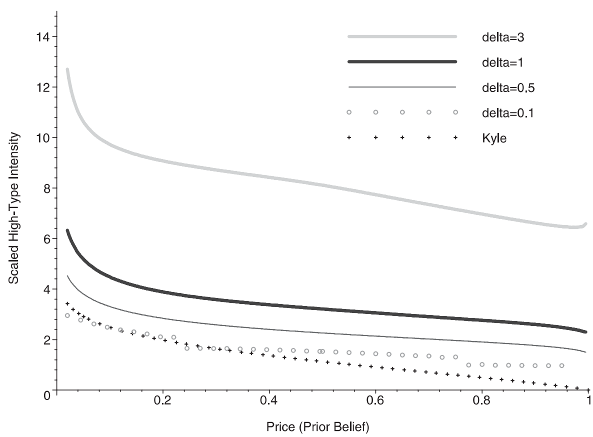

FIGURE 5.—Convergence of order intensities. Assuming $r = \sigma = 1$ and $2\beta\delta^2 = 1$, this shows the convergence of $\delta(\theta_{HB} - \theta_{HS})$ as δ decreases. The bottom curve (labeled "Kyle") is a plot of the equilibrium order rate θ_H^* of the high-type informed trader in our version of the Kyle model.

[6]As we explain in Appendix B, the θ's are estimated with significantly lower accuracy than are the value functions or ask and bid prices when bluffing is optimal. This accounts for the slightly anomalous behavior in Figure 5 of the plot for $\delta = .1$ for p below about .25 and p above about .75.

4. BLUFFING

We say that the informed trader is "bluffing" if he knowingly sells an undervalued asset or buys an overvalued asset. The benefit for the high type of selling is that it increases the market makers' belief that the asset value may actually be low, allowing the trader to buy later at a lower price. According to part (b) of the Corollary, the high type bluffs when p is large and δ is small. This is supported by our numerical results. Figure 6 plots

$$V_\delta(p) + (1 - b_\delta(p))\delta - V(b_\delta(p)). \tag{4.1}$$

The inequality (2.11) states that this quantity must be nonnegative. When it is positive, bluffing is strictly suboptimal. When it is zero, the high-type informed trader is willing to bluff. Figure 6 confirms part (b) of the Corollary, because it shows that bluffing is optimal for the high type for small values of δ and large values of p.

It seems reasonable that, if bluffing occurs at all, it will occur for the high type when p is large and δ is small, because the cost of bluffing is the immediate loss $-(1-b_\delta(p))\delta$, which is small when p is large (because $b_\delta(p)$ is then close to 1) and δ is small. On the other hand, the benefit of bluffing is the change in market makers' beliefs, and, according to part (b) of Theorem 3, the change in beliefs $b_\delta(p) - p$ is small when δ is small. To understand which of these factors is more important, we compute the ratio $(1 - b_\delta(p))\delta/(p - b_\delta(p))$. This is a type of "cost-benefit" ratio of bluffing for the high type. According to Theorem 3, this ratio is approximately $(1 - p)/\lambda^*(p)$ for small δ. As far as we know, there is no particular value of this cost-benefit ratio that indicates when bluffing will

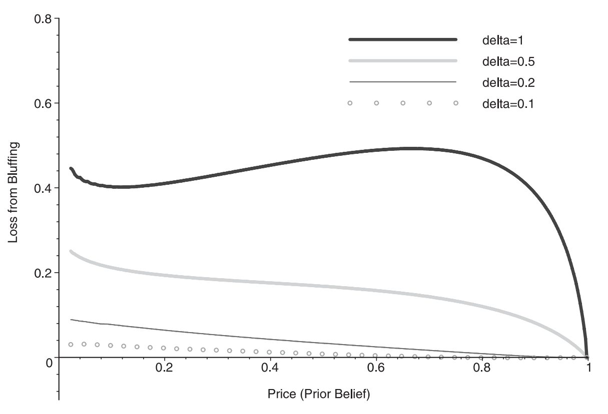

FIGURE 6.—The bluffing inequality. This plots inequality (2.11), which is the condition that bluffing cannot be better than optimal for the high-type informed trader. The value must be nonnegative, and it must be zero if the high type bluffs with positive probability. The plot shows there is no bluffing when $\delta = 1$ or $\delta = .5$, but bluffing is optimal for the high type when δ is small and p is large.

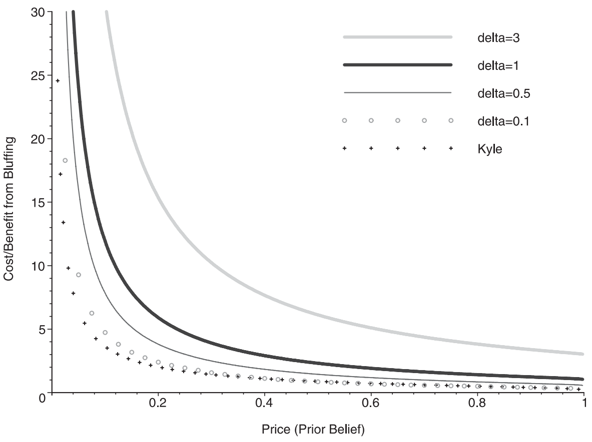

FIGURE 7.—Cost-benefit of bluffing. Assuming $r = \sigma = 1$ and $2\beta\delta^2 = 1$, this plots $(1 - b_\delta)\delta/(p - b_\delta)$, which is the "cost-benefit" ratio of bluffing for the high type. As δ decreases, this converges to the curve labeled "Kyle," which is a plot of $(1 - p)/\lambda^*(p)$. The figure shows that the relative cost of bluffing for the high type is lower as δ decreases and is lower as p increases.

be optimal. However, the numerical values of the ratio, shown in Figure 7, are consistent with the Corollary: the ratio is smaller when δ is smaller and it is smaller when p is large. Symmetrically, the cost-benefit ratio is likewise smaller for the low type when δ is smaller, and it is smaller for the low type when p is small.

The existence of bluffing may seem odd, but actually the same possibility exists in Kyle's (1985) model. Kyle's conclusion is that the informed trader buys/sells in proportion to the mispricing of the asset, but buying at rate f can be accomplished by buying at rate g and selling at rate h, as long as $f = g - h$. The informed trader is indifferent between these two alternatives and the market makers cannot distinguish them, because they see only the net orders. Hence, the conclusion of a unique linear equilibrium in Kyle (1985) should really be qualified: it is only the *net* buy rate of the informed trader that is unique; the actual buy and sell rates are indeterminate, except for the fact that the difference between them must equal the unique net buy rate. Seen in this light, it is perhaps less surprising that it is the expected *net* buy order of the high type in the Glosten–Milgrom model that converges to θ_H^* and the expected *net* sell order of the low type that converges to θ_L^*.

5. CONCLUSION

The Kyle (1985) and Glosten–Milgrom (1985) models are models of the same phenomenon, elucidated by Bagehot (1971): trades move prices because there is a possibility that the trader is better informed than the market at large. Glosten and Milgrom focus on the behavior of transactions prices, showing that if market makers are competitive and risk neutral, then transactions prices will be a martingale relative to their information (there is no negative serial correlation due to "bid-ask bounce") and prices in the long run will reflect the information of better informed traders. Kyle focuses on the "depth" or "liquidity" of the market, which depends on the amount of private information relative to the volume of uninformed trading. To derive the equilibrium depth, Kyle solves for the equilibrium strategy of an informed trader. In contrast, Glosten and Milgrom assume that the arrival rates of informed and uninformed traders are determined exogenously. However, Kyle makes the simplifying assumption that the market is organized as a series of batch auctions, which is not characteristic of most markets. The contribution of this paper is to show the consistency of the Kyle and Glosten–Milgrom models. We solved for the equilibrium strategy of an informed trader in a version of the Glosten–Milgrom model, and we showed that the equilibrium of the Glosten–Milgrom model is approximately the same as the equilibrium of the Kyle model, when the trade size is small and uninformed trades arrive frequently.

Our version of the Glosten–Milgrom model is less tractable than the continuous-time Kyle model, and we were only able to solve for the equilibrium numerically. Our results provide justification for using the more tractable continuous-time Kyle model, not just as an approximation to a discrete-time batch-auction market, but also as an approximation to a market in which individual trades are observed and executed by market makers, as in the Glosten–Milgrom model and as most markets are actually organized. Conversely, our results provide justification for using the Glosten–Milgrom model with exogenously imposed probabilistic arrival of informed and uninformed trades, because we showed that probabilistic arrival is consistent with equilibrium behavior when an informed trader is allowed to optimize his trading times.

John M. Olin School of Business, Washington University in St. Louis, St. Louis, MO 63130, U.S.A.; back@olin.wustl.edu

and

David Eccles School of Business, University of Utah, Salt Lake City, UT 84112, U.S.A.; finsb@business.utah.edu.

Manuscript received January, 2002; final revision received June, 2003.

APPENDIX A: PROOFS

PROOF OF THEOREM 1: For convenience, we will write $v = \tilde{v}$. First we will verify the filtering equations (1.6)–(1.8). Consider the system of stochastic differential equations

$$dp_t = \lambda(p_t)\{dY_t - \phi(p_t)\,dt\},$$

$$dY_t = [v\theta_H(p_t) - (1-v)\theta_L(p_t)]\,dt + dZ_t,$$

where λ is defined in (1.8) and ϕ in (1.6). We consider this system on $\{t : 0 < p_t < 1\}$. Set $\hat{v}_t = E[v|\mathcal{F}_t^Y]$. We need to show that $p_t = \hat{v}_t$.

Define

$$h_t = \frac{v\theta_H(p_t) - (1-v)\theta_L(p_t)}{\sigma}.$$

Set $z_t = Z_t/\sigma$ and $y_t = Y_t/\sigma$. Then, substituting for λ and ϕ from (1.8) and (1.6), we can write the system of stochastic differential equations as

$$\text{(A.1)} \qquad dp_t = \frac{p_t(1-p_t)[\theta_H(p_t) + \theta_L(p_t)]}{\sigma}\left\{dy_t - \frac{p_t\theta_H(p_t) - (1-p_t)\theta_L(p_t)}{\sigma}\,dt\right\},$$

$$\text{(A.2)} \qquad dy_t = h_t\,dt + dz_t.$$

Note z is a Wiener process.

Define $g_t = vh_t$. Because v is zero or one, we have $v^2 = v$ and therefore

$$g_t = \frac{v\theta_H(p_t)}{\sigma}.$$

Set $\hat{h}_t = E[h_t|\mathcal{F}_t^Y]$ and $\hat{g}_t = E[g_t|\mathcal{F}_t^Y]$. From the Fujisaki–Kallianpur–Kunita Theorem (Rogers and Williams (2000, Section VI.8)), we have

$$d\hat{v}_t = [\hat{g}_t - \hat{v}_t\hat{h}_t]\{dy_t - \hat{h}_t\,dt\}.$$

Directly from the definitions, we have

$$\hat{h}_t = \frac{\hat{v}_t\theta_H(p_t) - (1-\hat{v}_t)\theta_L(p_t)}{\sigma},$$

$$\hat{g}_t = \frac{\hat{v}_t\theta_H(p_t)}{\sigma},$$

$$\hat{g}_t - \hat{v}_t\hat{h}_t = \frac{\hat{v}_t(1-\hat{v}_t)[\theta_H(p_t) + \theta_L(p_t)]}{\sigma}.$$

Therefore,

$$\text{(A.3)} \qquad d\hat{v}_t = \frac{\hat{v}_t(1-\hat{v}_t)[\theta_H(p_t) + \theta_L(p_t)]}{\sigma}\left\{dy_t - \frac{\hat{v}_t\theta_H(p_t) - (1-\hat{v}_t)\theta_L(p_t)}{\sigma}\,dt\right\}.$$

From the definition of a conditional expectation of a 0–1 random variable, 0 and 1 must be absorbing boundaries for $\hat{v}$.

Now consider $\theta_H = \theta_H^*$ defined in (1.11) and $\theta_L = \theta_L^*$ defined in (1.12). These functions are continuously differentiable, hence locally Lipschitz and bounded on each compact subset of $(0, 1)$. This implies there is a unique solution of the system (A.1)–(A.3) up to any finite explosion time (Protter (1990, p. 199)). But the local boundedness of the θ's implies that y can explode only at a hitting time of the boundaries for p. Hence there is a unique solution of (A.1)–(A.3) on $\{t : 0 < p_t < 1\}$. But if $(p_t, y_t, \hat{v}_t)$ is a solution of (A.1)–(A.3), then clearly (p_t, y_t, p_t) is also a solution. Moreover, the boundaries are absorbing for both p and $\hat{v}$; therefore, $p_t = \hat{v}_t$ for all t.

The definition of θ_H^* and θ_L^* imply (1.6) and (1.8) with $\phi^* = 0$, so we conclude that the price dynamics stated in the theorem, $dp = \lambda^*\,dY$, imply $p_t = E[v|(Y_s)_{s\le t}]$ for $t < \tau$ as required. It remains to show that θ_H^* and θ_L^* are optimal. We will do this using the value functions given in (1.13) and (1.14) and simultaneously show that they are the true value functions.

We will prove optimality for θ_H^*. The proof of optimality for θ_L^* is identical. So assume $v = 1$. Let $V(p)$ denote the value given in (1.13); i.e.,

$$\text{(A.4)} \qquad V(p) = \int_p^1 \frac{1-a}{\lambda^*(a)}\,da.$$

We will write $\lambda = \lambda^*$ through the remainder of the proof. Obviously, (A.4) implies (1.15), the first part of the Bellman equation. Furthermore, the formula for λ implies (1.21), so we have

$$V(p) = \int_p^1 \frac{1-a}{\lambda(a)}\,da = -\frac{\sigma^2}{2r}\int_p^1 (1-a)\lambda''(a)\,da.$$

Integrating by parts and using the fact that $\lambda(1) = 0$ gives

$$\int_p^1 (1-a)\lambda''(a)\,da = -(1-p)\lambda'(p) - \lambda(p),$$

so

$$rV = \frac{1}{2}\sigma^2[\lambda + (1-p)\lambda']. \tag{A.5}$$

When we combine (A.5) with (1.15) we obtain (1.16), the second part of the Bellman equation. Furthermore, differentiating both sides of (1.10) and using the fact that $\lim_{p\to 0}\lambda(p) = 0$, we deduce that $\lim_{p\to 0}\lambda'(p) = \infty$. Therefore, (A.5) implies $\lim_{p\to 0} V(p) = \infty$.

Let X be an arbitrary admissible strategy of the form $dX = \theta_t\,dt$. The price dynamics $dp_t = \lambda(p_t)\,dY_t$ hold only for $t < \tau$ (because at time τ the price jumps from $p_{\tau-}$ to v) and for t less than the first date, if any, at which p_t hits the boundary. Consider a different process defined by $\hat{p}_0 = p_0$ and $d\hat{p}_t = \lambda(\hat{p}_t)[\theta_t\,dt + dZ_t]$ for all $t > 0$ and $0 < \hat{p}_t < 1$. We maintain the condition that 0 and 1 are absorbing states. We have of course that $\hat{p}_t = p_t$ for $t < \tau$.

By the definition of an admissible strategy,

$$E\int_0^\tau (1-p_u)\theta_u^-\,du < \infty.$$

Since τ is exponentially distributed with parameter r and independent of $\hat{p}$ and θ and since $p_t = \hat{p}_t$ for $t < \tau$, we have

$$E\int_0^\tau (1-p_u)\theta_u^-\,du = E\int_0^\infty e^{-ru}(1-\hat{p}_u)\theta_u^-\,du.$$

In particular,

$$\int_0^\infty e^{-ru}(1-\hat{p}_u)\theta_u^-\,du$$

exists and is finite a.s. This implies that

$$\int_0^\infty e^{-ru}(1-\hat{p}_u)\theta_u\,du \tag{A.6}$$

is well defined (though possibly equal to $+\infty$).

Let T denote the first time at which $\hat{p}$ hits the boundary; i.e., $T = \inf\{t \mid \hat{p}_t = 0 \text{ or } 1\}$. As usual, let $t\wedge T$ denote $\min\{t, T\}$. Applying Itô's lemma to the function $e^{-rt}V(\hat{p}_t)$ gives

$$\begin{aligned}
&e^{-r(t\wedge T)}V(\hat{p}_{t\wedge T}) - V(p_0)\\
&\quad= -\int_0^{t\wedge T} re^{-ru}V\,du + \int_0^{t\wedge T} e^{-ru}V'\,d\hat{p}_u + \frac{1}{2}\sigma^2\int_0^{t\wedge T} e^{-ru}\lambda^2V''\,du\\
&\quad= \int_0^{t\wedge T} e^{-ru}\left(-rV + \lambda\theta V' + \frac{1}{2}\sigma^2\lambda^2V''\right)du + \int_0^{t\wedge T} e^{-ru}\lambda V'\,dZ_u.
\end{aligned}$$

Substituting $V'(\hat{p}) = (\hat{p}-1)/\lambda(\hat{p})$ from (1.15) and

$$rV(\hat{p}) = \sigma^2\lambda(\hat{p})^2V''(\hat{p})/2$$

from (1.16) gives

$$e^{-r(t\wedge T)}V(\hat{p}_{t\wedge T}) - V(p_0) = -\int_0^{t\wedge T} e^{-ru}(1-\hat{p}_u)\theta_u\,du - \int_0^{t\wedge T} e^{-ru}(1-\hat{p}_u)\,dZ_u. \tag{A.7}$$

In conjunction with the nonnegativity of V this implies

$$V(p_0) \geq \int_0^{t\wedge T} e^{-ru}(1-\hat{p}_u)\theta_u\,du + \int_0^{t\wedge T} e^{-ru}(1-\hat{p}_u)\,dZ_u. \tag{A.8}$$

The second-term on the right-hand side of (A.8) almost surely has a finite limit as $t \to \infty$, because it is an L^2-bounded martingale (in fact, the L^2 norms are bounded by $1/2r$). We have already seen that the first term has a well-defined limit. Hence, we have

$$V(p_0) \geq \int_0^{T} e^{-ru}(1-\hat{p}_u)\theta_u\,du + \int_0^{T} e^{-ru}(1-\hat{p}_u)\,dZ_u, \tag{A.9}$$

and both integrals must be finite.

Consider the states of the world, if any, in which $T < \infty$. We argue that we must have $\hat{p}_T = 1$. If $\hat{p}_T = 0$, then the left-hand side of (A.7) is infinite, because $V(0) = \infty$. For the right-hand side to be infinite, we must have

$$\int_0^{T} e^{-ru}(1-\hat{p}_u)\theta_u\,du = -\infty,$$

which implies that $\int_0^T (1-\hat{p}_u)\theta_u\,du = -\infty$, which is precluded by admissibility. Therefore, $\hat{p}_T = 1$. Because $\hat{p}_u = \hat{p}_T = 1$ for $u \geq T$, (A.9) implies in this case that

$$V(p_0) \geq \int_0^{\infty} e^{-ru}(1-\hat{p}_u)\theta_u\,du + \int_0^{\infty} e^{-ru}(1-\hat{p}_u)\,dZ_u. \tag{A.10}$$

Of course, if $T = \infty$, then (A.9) is the same as (A.10), so we conclude that (A.10) holds almost surely.

Taking expectations throughout (A.10) yields

$$V(p_0) \geq E\int_0^{\infty} e^{-ru}(1-\hat{p}_u)\theta_u\,du,$$

because the stochastic integral, being an L^2-bounded martingale, is closed on the right by its limit. Using again the fact that τ is exponentially distributed with parameter r and independent of $\hat{p}$ and θ and the fact that $p_t = \hat{p}_t$ for $t < \tau$, we have

$$E\int_0^{\infty} e^{-ru}(1-\hat{p}_u)\theta_u\,du = E\int_0^{\tau}(1-p_u)\theta_u\,du.$$

So, we conclude that $V(p_0)$ is an upper bound on the expected profits.

We now consider the strategy $\theta = \theta_H^*$ $(= \theta_H)$ specified in the theorem. This strategy is certainly admissible because there are no losses for the high type when he never sells. We will show that the upper bound $V(p_0)$ is attained by this strategy. The key is to show that $\hat{p}_t \to 1$, implying that $e^{-rt}V(\hat{p}_t) \to 0$. Observe that $\hat{p}$ is a submartingale on the informed trader's filtration, because $d\hat{p}_t = \lambda(\hat{p}_t)[\theta_H(\hat{p}_t)\,dt + dZ_t]$ and $\theta_H > 0$. Because the submartingale $\hat{p}$ is bounded, it converges a.s. to some p_∞. We have

$$\begin{aligned} Ep_\infty &= \lim_{t\to\infty} E\hat{p}_t = p_0 + \lim_{t\to\infty} E\int_0^t \lambda(\hat{p}_u)\theta_H(\hat{p}_u)\,du \\ &= p_0 + \int_0^\infty E[\lambda(\hat{p}_u)\theta_H(\hat{p}_u)]\,du. \end{aligned}$$

The first equality is due to the bounded convergence theorem. The second equality is due to the fact that $E\int_0^t \lambda(\hat{p}_u)\,dZ_u = 0$ for all t, which follows from the boundedness of λ. The third equality is due to Fubini's theorem and the monotone convergence theorem. The finiteness of the improper integral on the right-hand side and the nonnegativity of the integrand implies $E[\lambda(\hat{p}_t)\theta_H(\hat{p}_t)] \to 0$. Now Fatou's lemma gives us $E[\lambda(p_\infty)\theta_H(p_\infty)] = 0$, which implies $p_\infty \in \{0, 1\}$ a.s.

We have

$$p_t I_{[t<\tau]} = E\big[vI_{[t<\tau]}|\mathcal{F}_t^y\big],$$

and it is easy to show that

$$p_t(1-p_t)I_{[t<\tau]} = E\big[(v-p_t)^2 I_{[t<\tau]}|\mathcal{F}_t^y\big].$$

Since $p = \hat{p}$ on the event $[t < \tau]$, this implies

$$\hat{p}_t(1-\hat{p}_t)I_{[t<\tau]} = E\big[(v-\hat{p}_t)^2 I_{[t<\tau]}|\mathcal{F}_t^y\big].$$

Now by iterated expectations and the independence of τ and $\hat{p}$ we obtain $E[\hat{p}_t(1-\hat{p}_t)] = E[(v-\hat{p}_t)^2]$. Since $p_\infty \in \{0,1\}$, $\hat{p}_t(1-\hat{p}_t) \to 0$ a.s. The bounded convergence theorem implies $E[(v-\hat{p}_t)^2] \to 0$. Therefore $\hat{p}_t \to v$ in the L^2 norm. This implies that there is a subsequence that converges to v a.s. However, every subsequence converges a.s. to p_∞. Therefore, $p_\infty = v$, and in the case at hand $p_\infty = 1$.

Now we reconsider the argument leading to (A.10). The inequality in (A.10) is due to the possibility that $\limsup_{t\to\infty} e^{-rt}V(\hat{p}_t) > 0$. However for this strategy we have $e^{-rt}V(\hat{p}_t) \to 0$, so we have equality in (A.10) and taking expectations gives

$$V(p_0) = E\int_0^\infty e^{-ru}(1-\hat{p}_u)\theta_H(\hat{p}_u)\,du = E\int_0^\tau (1-p_u)\theta_H(p_u)\,du.$$

This establishes optimality and that V is the value function. *Q.E.D.*

Proof of Theorem 2: Without loss of generality we will take $\delta = 1$. It follows from (2.5)–(2.8) that $p_t = E[v|\mathcal{F}_t^y]$. We will establish the optimality of the informed trading strategy.

Consider the high-type informed trader; i.e., $\tilde{v} = 1$. Let $x = x^+ - x^-$ be an arbitrary admissible strategy. In analogy to the proof of Theorem 1, let $\hat{p}$ denote the solution of $\hat{p}_0 = p_0$ and

$$d\hat{p}_t = f(\hat{p}_t)\,dt + [a(\hat{p}_{t-}) - p_{t-}]\,dy_t^+ + [b(\hat{p}_{t-}) - p_{t-}]\,dy_t^-,$$

so $\hat{p}_t = p_t$ for $t < \tau$. Let $T = \inf\{t \mid \hat{p}_t = 0 \text{ or } 1\}$.

By the definition of an admissible strategy,

$$E\int_0^\tau [b(p_{u-}) - 1]\,dx^- > -\infty.$$

Since τ is exponentially distributed with parameter r and independent of $\hat{p}$ and x, we have

$$E\int_0^\tau [b(p_{u-}) - 1]\,dx^- = E\int_0^\infty e^{-ru}[b(\hat{p}_{u-}) - 1]\,dx^-. \tag{A.11}$$

In particular,

$$\int_0^\infty e^{-ru}[b(\hat{p}_{u-}) - 1]\,dx^-$$

exists and is finite a.s. This implies that

$$\int_0^\infty e^{-ru}[1 - a(\hat{p}_{u-})]\,dx^+ + \int_0^\infty e^{-ru}[b(\hat{p}_{u-}) - 1]\,dx^- \tag{A.12}$$

is well defined (though possibly equal to $+\infty$).

Using the chain rule for Lebesgue–Stieltjes integrals and the law of motion for $d\hat{p}_t$ and substituting from (2.13), we have

$$
\begin{aligned}
\text{(A.13)} \qquad e^{-r(t\wedge T)}V(\hat{p}_{t\wedge T}) &= V(p_0) + \int_0^{t\wedge T} d(e^{-ru}V(\hat{p}_u)) \\
&= V(p_0) - r\int_0^{t\wedge T} e^{-ru}V(\hat{p}_{u-})\,du + \int_0^{t\wedge T} e^{-ru}\,dV(\hat{p}_u) \\
&= \int_0^{t\wedge T} e^{-ru}[V(a(\hat{p}_{u-})) - V(\hat{p}_{u-})]\,dx^+ \\
&\quad + \int_0^{t\wedge T} e^{-ru}[V(b(\hat{p}_{u-})) - V(\hat{p}_{u-})]\,dx^- \\
&\quad + \int_0^{t\wedge T} e^{-ru}[V(a(\hat{p}_{u-})) - V(\hat{p}_{u-})]\{dz^+ - \beta\,dt\} \\
&\quad + \int_0^{t\wedge T} e^{-ru}[V(b(\hat{p}_{u-})) - V(\hat{p}_{u-})]\{dz^- - \beta\,dt\}.
\end{aligned}
$$

The last two stochastic integrals on the right-hand side are L^2-bounded martingales, the boundedness being due to the bounds on $V(a) - V(p)$ and $V(b) - V(p)$ given by (2.9) and (2.11). Hence they converge almost surely to finite limits. We will denote the limit random variables by M and N. Rearranging, substituting for $V(a) - V(p)$ from (2.9), and taking limits therefore gives

$$
\begin{aligned}
\limsup_{t\to\infty} \bigg\{ &\int_0^{t\wedge T} e^{-ru}[1 - a(\hat{p}_{u-})]\,dx^+ \\
&+ \int_0^{t\wedge T} e^{-ru}[V(\hat{p}_{u-}) - V(b(\hat{p}_{u-}))]\,dx^- \bigg\} \\
= V(p_0) &+ M + N - \lim_{t\to\infty} e^{-r(t\wedge T)}V(\hat{p}_{t\wedge T}),
\end{aligned}
$$

assuming $\lim_{t\to\infty} e^{-r(t\wedge T)}V(\hat{p}_{t\wedge T})$ exists. In any case, the nonnegativity of V implies that the left-hand side is dominated by $V(p_0) + M + N$. In fact the condition $b - 1 \le V(p) - V(b)$ from (2.11) and the existence of the limit (A.12) implies

$$
\text{(A.14)} \qquad \int_0^T e^{-ru}[1 - a(\hat{p}_{u-})]\,dx^+ + \int_0^T e^{-ru}[b(\hat{p}_{u-}) - 1]\,dx^- \le V(p_0) + M + N,
$$

and both integrals on the left-hand side must be finite.

Consider the states of the world, if any, in which $T < \infty$. We claim that we must have $\hat{p}_T = 1$. If $\hat{p}_T = 0$, then the left-hand side of (A.13) equals $+\infty$ at $t = T$ by the assumed boundary condition $V(0) = \infty$. For the right-hand side to be infinite, we must have

$$
\int_0^T e^{-ru}[V(b(\hat{p}_{u-})) - V(\hat{p}_{u-})]\,dx^- = \infty,
$$

which, given that $1 - b \ge V(b) - V(p)$, implies

$$
\int_0^T e^{-ru}[1 - b(\hat{p}_{u-})]\,dx^- = \infty.
$$

However, this contradicts admissibility. Hence, we conclude that $\hat{p}_T = 1$. The formulas (2.5) and (2.6) imply $a(1) = b(1) = 1$. Because 1 is an absorbing point for $\hat{p}$, we have $a(\hat{p}_{u-}) = b(\hat{p}_{u-}) = 1$ for $u \geq T$. Therefore (A.14) implies

$$\text{(A.15)} \qquad \int_0^\infty e^{-ru}[1 - a(\hat{p}_{u-})]\,dx^+ + \int_0^\infty e^{-ru}[b(\hat{p}_{u-}) - 1]\,dx^- \leq V(p_0) + M + N.$$

If $T = \infty$, (A.14) is the same as (A.15), so we conclude that (A.15) holds almost surely.

Since the limit of an L^2-bounded martingale closes it on the right, taking expectations throughout (A.15) yields

$$E\int_0^\infty e^{-ru}[1 - a(\hat{p}_{u-})]\,dx^+ + E\int_0^\infty e^{-ru}[b(\hat{p}_{u-}) - 1]\,dx^- \leq V(p_0).$$

Using again the fact that τ is exponentially distributed with parameter r and independent of $\hat{p}$ and x and the fact that $p_t = \hat{p}_t$ for $t < \tau$, we conclude that

$$\text{(A.16)} \qquad E\int_0^\tau [1 - a(p_{u-})]\,dx^+ + E\int_0^\tau [b(p_{u-}) - 1]\,dx^- \leq V(p_0).$$

Therefore, $V(p_0)$ is an upper bound on the expected profits.

We will now show that the upper bound is attained by the strategy specified in the theorem. First we will show that under this strategy we have $\hat{p}_t \to 1$ a.s., implying $e^{-rt}V(\hat{p}_t) \to 0$ a.s., by virtue of the boundary condition $V(1) = 0$. The conditions (2.5)–(2.8) in conjunction with the conditions $0 < b < p < a < 1$, $\theta_{HS} < \theta_{LS}$, and $\theta_{LB} < \theta_{HB}$ imply that

$$k(p) \equiv f(p) + (a(p) - p)[\beta + \theta_{HB}] + (b(p) - p)[\beta + \theta_{HS}] > 0.$$

This is the expected change in $\hat{p}$ per unit of time conditional on $\tilde{v} = 1$. In fact, for $s < t$,

$$E[\hat{p}_{t\wedge T} - \hat{p}_{s\wedge T} | \mathcal{F}_s^y \vee \{\tilde{v} = 1\}] = E\left[\int_{s\wedge T}^{t\wedge T} k(\hat{p}_{u-})\,du \middle| \mathcal{F}_s^y \vee \{\tilde{v} = 1\}\right] \geq 0,$$

where $T \equiv \inf\{u | \hat{p}_{u-} \in \{0, 1\}\}$. This implies that $\hat{p}_{t\wedge T}$ is a submartingale. Since it is bounded by one, it converges almost surely to an integrable limit p_∞ (see Karatzas and Shreve (1988, p. 17)). In particular, if we take $s = 0$ and let t go to infinity, we get

$$\text{(A.17)} \qquad E\int_0^T k(\hat{p}_{u-})\,du = E[p_\infty | \tilde{v} = 1] - p_0 < \infty.$$

Since the expectation on the left-hand side is finite, the integral $\int_0^T k(\hat{p}_{u-})\,du$ is finite a.s.

If $T = \infty$, we conclude from the continuity and strict positivity of k that $\hat{p}_t \to p_\infty \in \{p : k(p) = 0\} = \{0, 1\}$ a.s. If $T < \infty$, then $\hat{p}_t = \hat{p}_T$ for all $t \geq T$ because 0 and 1 are absorbing, so in that case as well we have $\hat{p}_t \to p_\infty \in \{0, 1\}$ a.s.

Similarly, we can show that, conditioned on $\tilde{v} = 0$, $\hat{p}_{t\wedge T}$ is a positive supermartingale and $\hat{p}_t \to p_\infty \in \{0, 1\}$ a.s. The proof that $p_\infty = \tilde{v}$ can now be completed in the same way as for Theorem 1.

Now we reconsider the argument leading to (A.15). The inequality in (A.15) is due to the possibility that $\limsup_{t\to\infty} e^{-rt}V(\hat{p}_t) > 0$ and the possibility that $b - 1 < V(p) - V(b)$. However, for this strategy we have $e^{-rt}V(\hat{p}_t) \to 0$ and we have $dx^- = 1$ only when $b - 1 = V(p) - V(b)$. Therefore (A.15) holds with equality and taking expectations throughout yields equality in (A.16), establishing that this is an optimal strategy. *Q.E.D.*

PROOF OF THEOREM 3 AND THE COROLLARY: We adopt the hypothesis of the theorem. First we will show that $a_\delta(p) \to p$ and $b_\delta(p) \to p$ for each p. From (2.9), $V_\delta(p) - V_\delta(a_\delta(p)) \to 0$, so $V_\delta(a_\delta(p)) \to V(p)$. Because the sequence $\{a_\delta(p)\}$ is bounded, each subsequence has a further subsequence with some limit q. By uniform convergence, $V(q) = \lim V_\delta(a_\delta(p)) = V(p)$, which implies $q = p$ by the strict monotonicity of V. Because this is true for each subsequence, the sequence $\{a_\delta(p)\}$ must converge to p. The same argument using the convergence of J_δ to J and equation (2.10) shows that $b_\delta(p) \to p$.

Now we will show that $(a_\delta(p) - p)/\delta$ and $(p - b_\delta(p))/\delta$ have finite limits. Applying the mean value theorem to V_δ, we have

$$V'_\delta(\xi) = \frac{V_\delta(a_\delta(p)) - V(p)}{a_\delta(p) - p}$$

for some ξ between p and $a_\delta(p)$. Substituting from (2.9), this implies

$$V'_\delta(\xi) = \frac{-(1 - a_\delta(p))\delta}{a_\delta(p) - p}.$$

Since $a_\delta(p) \to p$, we have $\xi \to p$ and, by C^1 convergence, $V'_\delta(\xi) \to V'(p)$. Therefore

$$V'(p) = \lim_{\delta\to 0} \frac{-(1 - a_\delta(p))\delta}{a_\delta(p) - p}.$$

The factor $1 - a_\delta(p)$ has the nonzero limit $1 - p$, so $(a_\delta(p) - p)/\delta$ must also have a finite limit, which we will denote by $A(p)$. Now we can write

$$V'(p) = \frac{-(1 - p)}{A(p)}. \tag{A.18}$$

Likewise, using (2.10), we can show that $(p - b_\delta(p))/\delta$ has a finite limit $B(p)$ and

$$J'(p) = \frac{p}{B(p)}. \tag{A.19}$$

Now we will show that the limits $A(p)$ and $B(p)$ are equal. Applying the mean value theorem to V_δ again and using (2.11), we have, for some ξ between $b_\delta(p)$ and p,

$$\begin{aligned} V'_\delta(\xi) &= \frac{V_\delta(p) - V_\delta(b_\delta(p))}{p - b_\delta(p)} \\ &\geq \frac{-(1 - b_\delta(p))\delta}{p - b_\delta(p)}. \end{aligned}$$

Taking limits we obtain

$$V'(p) \geq \frac{-(1 - p)}{B(p)}. \tag{A.20}$$

Likewise, we obtain from (2.12) that

$$J'(p) \leq \frac{p}{A(p)}. \tag{A.21}$$

We conclude from (A.18) and (A.20) that $A \geq B$ and from (A.19) and (A.21) that $A \leq B$, so we must have $A = B$.

Define $\lambda = A = B$. Expand $V_\delta(a)$ and $V_\delta(b)$ in (2.13) by exact second-order Taylor series expansions to obtain

$$\text{(A.22)} \qquad rV_\delta(p) = V_\delta'(p)[f_\delta + \beta(a_\delta + b_\delta - 2p)] + \frac{1}{2}\beta V_\delta''(\xi)(a_\delta - p)^2 + \frac{1}{2}\beta V_\delta''(\xi')(b_\delta - p)^2,$$

for some $p < \xi < a_\delta(p)$ and $b_\delta(p) < \xi' < p$. Substituting $\beta = \sigma^2/2\delta^2$, we obtain from the previous paragraph and the C^2 convergence that the last two terms each converge to

$$\frac{1}{4}\sigma^2\lambda^2 V''(p).$$

Set

$$\text{(A.23)} \qquad \kappa_\delta(p) = f_\delta(p) + \beta[a_\delta(p) + b_\delta(p) - 2p] = f_\delta(p) + \frac{\sigma^2[a_\delta(p) + b_\delta(p) - 2p]}{2\delta^2}.$$

It must have a limit because the other terms in (A.22) converge and the limit of $V_\delta'(p)$ is nonzero. Denoting the limit by $\kappa(p)$ and defining $\phi(p) = -\kappa(p)/\lambda(p)$, we have

$$\text{(A.24)} \qquad rV(p) = -\phi(p)\lambda(p)V'(p) + \frac{1}{2}\sigma^2\lambda^2 V''(p).$$

The exact same reasoning applied to (2.14) yields

$$\text{(A.25)} \qquad rJ(p) = -\phi(p)\lambda(p)J'(p) + \frac{1}{2}\sigma^2\lambda^2 J''(p).$$

The system of equations (A.18)–(A.19) and (A.24)–(A.25), with $A = B = \lambda$, is the system (1.15)–(1.18) that characterizes equilibrium in our version of the Kyle model. We now give the argument that we sketched in Section 1 leading to $\phi = 0$ and the differential equation (1.21). Note that from (A.18) and the fact that V is C^2, λ must be C^1. Differentiating (A.18) gives

$$V'' = \frac{\lambda + (1-p)\lambda'}{\lambda^2},$$

and substituting this and (A.18) into (A.24) gives

$$\text{(A.26)} \qquad rV = (1-p)\phi + \frac{1}{2}\sigma^2[\lambda + (1-p)\lambda']$$

$$\text{(A.27)} \qquad = \frac{1}{2}\sigma^2\lambda + (1-p)\left[\phi + \frac{1}{2}\sigma^2\lambda'\right].$$

This implies that the function $\phi + \sigma^2\lambda'/2$ must be C^1, because the other terms are C^1. Differentiating again and substituting for V' from (A.18) gives

$$\text{(A.28)} \qquad \frac{r(p-1)}{\lambda} = -\phi + (1-p)\frac{d}{dp}\left[\phi + \frac{1}{2}\sigma^2\lambda'\right].$$

Repeating this argument for J gives

$$\text{(A.29)} \qquad \frac{rp}{\lambda} = -\phi - p\frac{d}{dp}\left[\phi + \frac{1}{2}\sigma^2\lambda'\right].$$

Subtracting (A.29) from (A.28) yields

$$\frac{-r}{\lambda} = \frac{d}{dp}\left[\phi + \frac{1}{2}\sigma^2\lambda'\right].$$

Dividing by $-p$ in (A.29) also yields

$$\frac{-r}{\lambda} = \frac{\phi}{p} + \frac{d}{dp}\left[\phi + \frac{1}{2}\sigma^2\lambda'\right].$$

Comparing these last two equations, we conclude that $\phi = 0$. We already noted that $\phi + \sigma^2\lambda'/2$ must be C^1, so actually it must be the case that λ is C^2. Moreover, equations (A.28) and (A.29) are each equivalent to (1.21).

Since λ is nonnegative, the differential equation (1.21) implies λ is concave. The limits $\lim_{p\to 0}\lambda(p)$ and $\lim_{p\to 1}\lambda(p)$ are therefore well defined. The boundary condition $V(0) = \infty$ implies $\limsup_{p\to 0} V'(p) = \infty$. In view of (A.18) this implies $\lim_{p\to 0}\lambda(p) = 0$. Similarly, the boundary condition $J(1) = \infty$ and (A.19) imply $\lim_{p\to 1}\lambda(p) = 0$.

The uniqueness of the solution to the differential equation (1.21) with these boundary conditions implies that λ is the λ^* given by (1.10). This verifies part (b) of the theorem. Furthermore the boundary conditions $V(1) = J(0) = 0$ in conjunction with (A.18) and (A.19) imply that V and J are given by (1.13) and (1.14), which is part (a).

It remains to prove part (c) and the Corollary. Fix any $p \in (0, 1)$. Part (c) of the theorem is trivially true unless both $\theta_{LB}(p)/\beta \to 0$ and $\theta_{HS}(p)/\beta \to 0$, so assume that these conditions hold; i.e., adopt the hypothesis of the Corollary.

Consider the quantity $\kappa_\delta(p)$ defined in (A.23). We have shown that $\kappa_\delta(p) \to 0$ for each p. From the definition (2.7) of f and part (b) we have

$$\begin{aligned}\kappa_\delta &\equiv f + \beta(a_\delta - p) + \beta(b_\delta - p)\\ &= -[p\theta_{HB} + (1-p)\theta_{LB}]\delta\left(\frac{a_\delta - p}{\delta}\right) + [p\theta_{HS} + (1-p)\theta_{LS}]\delta\left(\frac{p - b_\delta}{\delta}\right)\\ &\approx \delta[p\theta_{HS} + (1-p)\theta_{LS} - p\theta_{HB} - (1-p)\theta_{LB}]\lambda^*.\end{aligned}$$

Therefore, $\kappa_\delta \to 0$ implies

$$\delta[p\theta_{HS} + (1-p)\theta_{LS} - p\theta_{HB} - (1-p)\theta_{LB}] \to 0. \tag{A.30}$$

The ask price can be written (see the derivation above (2.5)) as

$$a_\delta(p) = m_\delta + n_\delta p,$$

where

$$m_\delta = \frac{p\theta_{HB}}{p\theta_{HB} + (1-p)\theta_{LB} + \beta}$$

and

$$n_\delta = \frac{\beta}{p\theta_{HB} + (1-p)\theta_{LB} + \beta}.$$

The assumption $\theta_{LB}/\beta \to 0$ implies $m_\delta + n_\delta \to 1$ and the fact that $a_\delta \to p$ now leads to $n_\delta \to 1$ and $m_\delta \to 0$. From $n_\delta \to 1$, we conclude that

$$\frac{p\theta_{HB} + (1-p)\theta_{LB}}{\beta} \to 0. \tag{A.31}$$

Likewise, we can deduce from $\theta_{HS}/\beta \to 0$ and $b_\delta \to p$ that

$$\frac{p\theta_{HS} + (1-p)\theta_{LS}}{\beta} \to 0. \tag{A.32}$$

From (2.5), the definition of β, and (A.31), we have

$$\begin{aligned}\frac{a_\delta - p}{\delta} &= \frac{2\beta\delta(a_\delta - p)}{\sigma^2} \\ &= \frac{2p(1-p)}{\sigma^2}\frac{\delta(\theta_{HB} - \theta_{LB})}{1 + [p\theta_{HB} + (1-p)\theta_{LB}]/\beta} \\ &\approx \frac{2p(1-p)}{\sigma^2}\delta(\theta_{HB} - \theta_{LB}).\end{aligned}$$

Therefore, the result $(a_\delta - p)/\delta \to \lambda^*$ implies

$$\delta(\theta_{HB} - \theta_{LB}) \to \frac{\sigma^2\lambda^*}{2p(1-p)}. \tag{A.33}$$

Likewise, (2.6), (A.32), and the result that $(p - b_\delta)/\delta \to \lambda^*$ imply

$$\delta(\theta_{LS} - \theta_{HS}) \to \frac{\sigma^2\lambda^*}{2p(1-p)}. \tag{A.34}$$

Write (A.30) as

$$\delta p(\theta_{HS} - \theta_{LS}) + \delta p(\theta_{LB} - \theta_{HB}) + \delta(\theta_{LS} - \theta_{LB}) \to 0.$$

From (A.34) the first term converges to $-\sigma^2\lambda^*/(2(1-p))$, and from (A.33) the second term converges to the same thing. Therefore,

$$\delta(\theta_{LS} - \theta_{LB}) \to \frac{\sigma^2\lambda^*}{1-p}. \tag{A.35}$$

Likewise, (A.30), (A.33), and (A.34) imply

$$\delta(\theta_{HB} - \theta_{HS}) \to \frac{\sigma^2\lambda^*}{p}. \tag{A.36}$$

The limit in (A.35) is the definition (1.12) of θ_L^*, and the limit in (A.36) is the definition (1.11) of θ_H^*. This completes the proof of part (a) of the Corollary.

Now write (A.30) as

$$\delta(1-p)(\theta_{LS} - \theta_{HS}) + \delta p(\theta_{LB} - \theta_{HB}) + \delta(\theta_{HS} - \theta_{LB}) \to 0.$$

From (A.34) the first term converges to $\sigma^2\lambda^*/(2p)$, and from (A.33) the second term converges to $-\sigma^2\lambda^*/(2(1-p))$. Therefore,

$$\delta(\theta_{HS} - \theta_{LB}) \to \frac{(2p-1)\sigma^2\lambda^*}{2p(1-p)}.$$

Thus, for $p > 1/2$ and sufficiently small δ,

$$\delta\theta_{HS} \geq \delta(\theta_{HS} - \theta_{LB}) > 0,$$

and, for $p < 1/2$ and sufficiently small δ,

$$\delta\theta_{LB} \geq \delta(\theta_{LB} - \theta_{HS}) > 0,$$

which is part (b) of the Corollary. As noted before, part (b) of the Corollary implies part (c) of the theorem. *Q.E.D.*

APPENDIX B: NUMERICAL METHOD

We work with a grid of size n on $[0, 1]$. We start with a guess for the value function V of the high type at each point on the grid, taking $V(1) = 0$. We interpolate linearly to compute V at all other points, except for $p < 1/n$. For $p < 1/n$, we extrapolate V as follows. We choose a normalizing function $\xi(p)$ that is also infinite at 0 and fit a quadratic function to $V(p)/\xi(p)$ at $p = 1/n$, $2/n$, and $3/n$. We then calculate $V(p)$ for $p < 1/n$ as $(ap^2 + bp + cp)\xi(p)$. For the normalizing function, we use the value function (1.13) from the Kyle model. We also start with a guess $f = 0$ for the drift of p between orders, the equilibrium value of which is given in (2.7).

We work on the grid in the region $p \geq 1/2$ and use symmetry to define variables for $p < 1/2$. In accordance with part (b) of the Corollary, we conjecture (and confirm) that $\theta_{LB}(p) = 0$ for $p > 1/2$. Note that in this circumstance equations (2.5)–(2.7) imply

$$f(p) = g(p) + \frac{p[p - b(p)]}{b(p)} \theta_{HS}(p),$$

where

$$g(p) = \frac{\beta p[p - b(p)]}{b(p)} - \frac{\beta(1 - p)[a(p) - p]}{1 - a(p)}. \tag{B.1}$$

The algorithm is as follows.

Step 1: We set $J(p) = V(1 - p)$. We compute $a(p)$ on the grid for $p \geq 1/2$ from

$$V(p) = (1 - a(p))\delta + V(a(p)),$$

and we compute $b(p)$ from

$$J(p) = b(p)\delta + J(b(p)).$$

We compute $g(p)$ from (B.1).

Step 2: We check inequality (2.11). If it is satisfied, we set $\hat{f}(p) = g(p)$. If not, we set

$$\hat{f}(p) = (1 - \varepsilon_1)f(p) + \varepsilon_1 \frac{rV(p) - \beta\delta[a(p) - b(p)]}{V'(p)} \tag{B.2}$$

for a suitably chosen constant ε_1. To estimate $V'(p)$ in (B.4), we use a five-point approximation (Gerald and Wheatley (1999, p. 373)). Equation (B.2) is motivated by (2.13), which states that

$$f(p) = \frac{rV(p) - \beta[V(a(p)) + V(b(p)) - 2V(p)]}{V'(p)}, \tag{B.3}$$

which equals

$$\frac{rV(p) - \beta\delta[a(p) - b(p)]}{V'(p)},$$

when (2.9) holds and (2.11) holds as an equality. We then compute $\hat{V}$ and $\hat{J}$ by adding a small fraction of the errors in (2.13) and (2.14). Specifically, we set

$$\hat{V}(p) = V(p) + \varepsilon_2\{V'(p)\hat{f}(p) + \beta[V(a(p)) + V(b(p)) - 2V(p)] - rV(p)\}, \tag{B.4}$$

$$\hat{J}(p) = J(p) + \varepsilon_2\{J'(p)\hat{f}(p) + \beta[J(a(p)) + J(b(p)) - 2J(p)] - rJ(p)\}. \tag{B.5}$$

Equations (B.4)–(B.5) employ the method that Judd (1998, p. 166) describes as "extrapolation."

Step 3: For $p < 1/2$, define $\hat{V}(p) = \hat{J}(1 - p)$ and return to Step 1, with $V = \hat{V}$ and $f = \hat{f}$.

When this converges, we have V, J, a, b, and f. We compute the θ's from a, b, and f via (2.5)–(2.7).

Our updating equation for the vaue functions can be understood as a type of value iteration. To interpret it in this way, consider a discrete-time model with period length Δt in which an uninformed buy order arrives with probability $\beta\Delta t$, an uninformed sell order arrives with the same probability, the announcement arrives with probability $r\Delta t$, and no two of these events occur simultaneously. Assume that, if none of these events occurs, then p moves to $p + f(p)\Delta t$, which is consistent with (2.8). Assuming it is optimal to wait to trade and letting V denote the value function for the next period, the value function $\hat{V}$ for the current period will satisfy

$$\hat{V}(p) = \beta\Delta t[V(a(p)) + V(b(p))] + (1 - r\Delta t - 2\beta\Delta t)V(p + f(p)\Delta t). \tag{B.6}$$

If we take $\Delta t = \varepsilon_2$, approximate $V(p + f(p)\Delta t)$ as $V(p) + V'(p)f(p)\Delta t$, and ignore terms of order $(\Delta t)^2$, (B.6) is the same as (B.4). Value iteration is well known to converge in discounted dynamic programming problems (see, e.g., Judd (1998, p. 412)), and one would expect this discrete-time model to converge to our continuous-time model as $\Delta t \to 0$. However, we are actually solving an equilibrium problem, updating the functions a, b, and f in each step, so we cannot appeal to standard results for convergence. Nevertheless, the algorithm did converge.

We iterated until the change in V (the maximum change across the grid points) was sufficiently small. We enforced (2.9)–(2.10), so the equilibrium conditions we need to check at the end are the inequalities (2.11) and (2.12) and the differential–difference equations (2.13) and (2.14). For the values of $\delta = .5$ and above shown in the figures, the inequalities hold strictly, as Figure 6 shows. For $\delta = .2$ and below, the inequality (2.11) holds as an equality for large values of p and (2.12) holds as an equality for small values of p, with a maximum error in all cases on the order of 10^{-4}. In all cases, the error in (2.13) and (2.14) was less than 10^{-10} at each grid point.

We did numerous robustness checks. We started with different initial estimates for V, we used different grid sizes ($n = 100$, $n = 400$, and $n = 1000$), we used different methods of extrapolating V below $1/n$, and we tried both (B.4) and (B.6) (and the equation for $\hat{J}$ corresponding to (B.6)) as the updating equation. Provided the constants ε_1 and ε_2 were chosen well, the algorithm converged in all cases, and when it converged, it converged to the same limits.

Even though all the equilibrium conditions hold with a high order of accuracy, it appears from our plots that the θ's are not estimated very accurately when there is bluffing. This is probably an inevitable result of our estimation method, because we are estimating the θ's from f obtained from (B.3). The derivative $V'(p)$ is small near one, where bluffing occurs, and even if V were known exactly, small errors introduced by the numerical computation of this derivative in (B.3) would lead to relatively large errors in f and hence in the θ's.

REFERENCES

BACK, K. (1992): "Insider Trading in Continuous Time," *Review of Financial Studies*, 5, 387–409.

BAGEHOT, W. (PSEUD.) (1971): "The Only Game in Town," *Financial Analysts Journal*, 22, 12–14.

BLACK, F. (1990): "Bluffing," Working Paper, Goldman Sachs.

EASLEY, D., N. KIEFER, M. O'HARA, AND J. PAPERMAN (1996): "Liquidity, Information, and Infrequently Traded Stocks," *Journal of Finance*, 51, 1405–1436.

ETHIER, S., AND T. KURTZ (1986): *Markov Processes*. New York: John Wiley & Sons.

GERALD, C., AND P. WHEATLEY (1999): *Applied Numerical Analysis*. Reading, MA: Addison-Wesley.

GLOSTEN, L., AND P. MILGROM (1985): "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 13, 71–100.

HUDDART, S., J. HUGHES, AND C. LEVINE (2001): "Public Disclosure and Dissimulation of Insider Trades," *Econometrica*, 69, 665–681.

JUDD, K. (1998): *Numerical Methods in Economics*. Cambridge: MIT Press.

KARATZAS, I., AND S. SHREVE (1988): *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.

KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

PROTTER, P. (1990): *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.

ROGERS, L. C. G., AND D. WILLIAMS (2000): *Diffusions, Markov Processes and Martingales*, 2nd ed. Cambridge, MA: Cambridge University Press.