Dept. of Math./CMA University of Oslo
Pure Mathematics No 14
ISSN 0806–2439 August 2010

Strategic Insider Trading Equilibrium: A Filter Theory Approach.

Knut K. Aase[2,1], Terje Bjuland[2] and Bernt Øksendal[1,2]
Knut.Aase@nhh.no, Terje.Bjuland@nhh.no, oksendal@math.uio.no

24 August 2010

Abstract

The continuous-time version of Kyle's (1985) model of asset pricing with asymmetric information is studied, and generalized in various directions, i.e., by allowing time-varying liquidity trading, and by having weaker a priori assumptions on the model. This extension is made possible by the use of filtering theory. We derive the optimal trade for an insider and the corresponding price of the risky asset; the insider's trading intensity satisfies a deterministic integral equation, given perfect inside information.

Mathematics Subject Classification 2010: 60G35, 62M20, 93E10, 94Axx

Keywords: Insider trading, equilibrium, strategic trade, linear filter theory, innovation equation

[01]Centre of Mathematics for Applications (CMA), Dept of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N–0316 Oslo, Norway.
The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [228087]
[02]Norwegian School of Economics and Business Administration (NHH), Helleveien 30, N–5045 Bergen, Norway.

1 Introduction

We take as our starting point the seminal paper of Kyle (1985), where a model of asset pricing with asymmetric information is presented. Traders submit order quantities to risk-neutral market makers, who set prices competitively by taking the opposite position to clear the market. Excluding the market makers, the model has two kinds of traders: a single risk neutral informed trader and liquidity (noise) traders. The informed trader rationally anticipates the effects of his orders on the price, i.e., she acts non-competitively or strategically. In the presence of noise traders it is impossible for the market makers to exactly invert the price and infer the informed trader's signal. Thus markets are semi-strong, but not strong form efficient.

In this model the insider makes positive profits in equilibrium by exploiting his monopoly power optimally in a dynamic context. Noise trading provides camouflage which conceals his trading from market makers. An important issue is to demonstrate that this is possible in equilibrium without destabilizing prices.

Kyle's approach is to first study a one-period auction, then extend the analysis to a model in with auctions take place sequentially, and finally letting the time between the auctions go to zero, in which case a limiting model of continuous trading is obtained. Back (1992) formalize and extend the continuous-time version of the Kyle model, by i.a., the use of dynamic programming.

There is a rich literature on the one period model, as well as on discrete insider trading, e.g., Holden and Subrahmanyam (1992), Admati and Pfleiderer (1988), and others, all adding insights to this class of problems. Glosten and Milgrom (1985) present a different approach, containing similar results to Kyle. Before Kyle (1985) and Glosten and Milgrom (1985) there is also a huge literature on insider trading in which the insider acts competitively, e.g., Grossman and Stiglitz (1980).

The approach of this article is to study the continuous-time model directly, not as a limiting model of a sequence of auctions, and use the machinery of filtering theory in continuous-time to resolve the problem, in a more general setting with time-varying noise trading. There are also other generalizations that our approach can handle in addition to the ones already mentioned: One is that we do not assume that the final price p_T equals the signal $\tilde{v}$, but show that this is a consequence of our other model assumptions.

We are able to both find the price of the risky asset and solve the insider's problem in a direct way, leading to a deterministic integral equation for the insider's trading intensity $\beta(t)$ at time t, given his information set with perfect forward information, and correlated liquidity trade.

We solve the integral equation for the trading intensity $\beta(t)$ by transforming this equation to a non-linear, separable differential equation, which calls for a simple solution. This we compare to the solution of Kyle (1985) (and also Back (1992)). In the special case of time homogeneous noise trading we recover the Kyle-solution. For time-varying noise trading we get the result that the market depth is still a constant, and the expected (ex ante) profits of the insider depends on the average volatility process.

2 The Model

At date T there is to be a public release of information that will perfectly reveal the value of an asset; cf. fair value accounting. Trading in this asset and a risk-free asset with interest rate zero is assumed to occur continuously during the interval $[0, T]$. The information to be revealed at time T is represented as a signal $\tilde{v}$, a random variable which we interpret as the price at which the asset will trade after the release of information. This information is already possessed by a single insider at time zero. The unconditional distribution of $\tilde{v}$ is assumed to be *normal* with mean $\mu_{\tilde{v}}$ and variance $\sigma_{\tilde{v}}^2$.

In addition to the insider, there are liquidity traders, and risk neutral market makers. The liquidity traders are unable to correlate their orders to the insider's signal $\tilde{v}$. Thus the liquidity traders have random, price-inelastic demands. All orders are market orders and the net order flow is observed by all market makers. We denote by z_t the cumulative orders of liquidity traders through time t. The process z is assumed to be a Brownian motion with mean zero and variance rate σ_t^2, i.e., $dz_t = \sigma_t dB_t$, where $\sigma_t > 0$ is a deterministic continuously differentiable function on $[0, T]$, for a standard Brownian motion B defined on a probability space (Ω, P). As Kyle (1985) and Back (1992) we assume that B is independent of $\tilde{v}$. We let x_t be the cumulative orders of the informed trader, and define

$$y_t = x_t + z_t \qquad \text{for all } t \in [0, T] \tag{2.1}$$

as the total orders accumulated by time t.

Market makers only observe the process y, so they cannot distinguish between informed and uninformed trades. Let $\mathcal{F}_t^y = \sigma(y_s; s \leq t)$ be the information filtration of this process. Since the market makers are assumed to be perfectly competitive and risk neutral, they will set the price p_t at time t as follows

$$p_t = E[\tilde{v}|\mathcal{F}_t^y], \tag{2.2}$$

which we will call a *rational* pricing rule. The market makers, the insider and the liquidity traders all know the probability distribution of $\tilde{v}$.

We assume that the insider's portfolio is of the form

$$dx_t = (\tilde{v} - p_t)\beta_t dt, \qquad x_0 = 0, \tag{2.3}$$

where $\beta \geq 0$ is some deterministic function. Both assumptions are consistent with Kyle (1985).[1] The function β_t is called the *trading intensity* on the information advantage $(v - p_t)$ of the insider.

Denote the insider's wealth by w and the investment in the risk-free asset by b. The budget constraint of the insider can best be understood by considering a discrete time model. At time t the agent submits a market order $x_t - x_{t-1}$ and the price changes from p_{t-1} to p_t. The order is executed at price p_t, in other words, $x_t - x_{t-1}$ is submitted *before* p_t is set by the market makers. The investment in the risk-free asset changes by $b_t - b_{t-1} = -p_t(x_t - x_{t-1})$, i.e., buying stocks leads to reduced cash with exactly the same amount. Thus, the associated change in wealth is (which was pointed out by Back (1992))

$$b_t - b_{t-1} + x_t p_t - x_{t-1} p_{t-1} = x_{t-1}(p_t - p_{t-1}). \tag{2.4}$$

In other words, the usual accounting identity for the wealth dynamics is of the same type as in the standard price-taking model, except for one important difference; while, in the rational expectations model, the number of stocks in the risky asset at time t is depending only on the information available at this time, so that both the processes x and p are adapted processes with respect to the same filtration, here the order x depends on information available only at time T for the market makers (and the noise traders). As a consequence we obtain the dynamic equation for the insider' wealth w_t as follows

$$w_t = w_0 + \int_0^t x_s dp_s \tag{2.5}$$

[1] The finite variation property of x is assumed by Kyle (1985), and an equilibrium where this is the case is found by Back (1992).

This is not well-defined as a stochastic integral in the traditional interpretation, since p_t is $\mathcal{F}_t^y$-adapted, and x_t is not. Thus it needs further explanation. However, since we assume that the strategy of the insider has the form (2.3) for some deterministic continuous function $\beta_t > 0$, then a natural interpretation of (2.5) is obtained by using *integration by parts*, as follows:

$$\begin{aligned} w_t &= w_0 + x_t p_t - \int_0^t p_s dx_s \\ &= w_0 + p_t \int_0^t (\tilde{v} - p_s)\beta_s ds - \int_0^t p_s(\tilde{v} - p_s)\beta_s ds \\ &= w_0 + \int_0^t (\tilde{v} - p_s)^2 \beta_s ds - \int_0^t (\tilde{v} - p_t)(\tilde{v} - p_s)\beta_s ds. \end{aligned} \tag{2.6}$$

Alternatively, one might obtain (2.6) by interpreting the stochastic integral in (2.5) as a *forward integral*. See Russo and Vallois (1993), Russo and Vallois (1995, 2000) for definitions and properties and Biagini and Øksendal (2005) for applications of forward integrals to finance.

The insider tries to find the trading intensity β_t which maximizes the expected terminal wealth

$$E[w_T] = w_0 + \int_0^T E[(\tilde{v} - p_s)^2]\beta_s ds - \int_0^T E[(\tilde{v} - p_T)(\tilde{v} - p_s)]\beta_s ds. \tag{2.7}$$

The dilemma for the insider is that an increased trading intensity at some time t will reveal more information about the value of $\tilde{v}$ to the market makers and hence induce a price p_t closer to $\tilde{v}$, which in turn implies a reduced insider information advantage.

Let us define the information filtration of the informed trader as $\mathcal{G}_t = \mathcal{F}_t^y \vee \sigma(\tilde{v})$. Thus the informed trader knows $\tilde{v}$ at time zero and observes y_t at each time t. Obviously the filtration $\mathcal{G}_t \supset \mathcal{F}_t^y$ and this extension is not of a trivial type, but a significant one. For example, there is information in $\mathcal{G}_t$ for any $t \in [0, T)$ that will only be revealed to the market makers at the future time T. The key point here is that from (2.3) the order x_t depends on $\tilde{v}$ which is not in $\mathcal{F}_t^y$. Since the insider knows the realization of $\tilde{v}$ at time 0, she has long-lived forward-looking information.

We can now formulate the problem mathematically:
The insider wants to solve
(2.8)

$$\max_{\beta} E[w_T] = w_0 + \max_{\beta}(\int_0^T E[(\tilde{v} - p_s)^2]\beta_s ds - \int_0^T E[(\tilde{v} - p_T)(\tilde{v} - p_s)]\beta_s ds).$$

subject to the price p satisfying the rational pricing rule (2.2), for all $t \in [0, T]$.

Usually the assumption is made that $\lim_{s \to T^-} p_t = p_T = \tilde{v}$ a.s., but as we will show below, this is a consequence of our other model assumptions, *provided that the insider trades optimally.* This result seems natural, ensuring that all information available has been incorporated in the price at the time T of the public release of the information. But note that if the insider does not trade optimally then this need not hold.

Since there is a tacit understanding that the price process p is continuous in this model, this result also means that the insider must trade continuously throughout the time interval $[0, T]$, and we can expect that the trading intensity β must be large as t approaches T in order for this condition to be satisfied. [2]

An *equilibrium* is a pair (p, x) such that p satisfies (2.2), given x, and x is an optimal trading strategy solving (2.8), given p. Moreover, we require that the *mean square error process* $S_t(\beta)$ satisfies

$$S_t(\beta) := E[(\tilde{v} - p_t)^2] > 0 \quad \text{for all } t \in [0, T). \tag{2.9}$$

Here $S_0(\beta) := S_0 := \sigma_{\tilde{v}}^2$. This assumption will be discussed and relaxed later.

We now have the following result:

Theorem 2.1. *The optimal trading intensity β_t of the insider is given by*

$$\beta_t = \frac{S_0^{1/2}(\int_0^T \sigma_s^2 ds)^{1/2}\sigma_t^2}{S_0 \int_t^T \sigma_s^2 ds}; \qquad t \in [0, T). \tag{2.10}$$

The corresponding optimal wealth of the insider is

$$J(\beta) = S_0^{1/2}\Big(\int_0^T \sigma_t^2 dt\Big)^{1/2}. \tag{2.11}$$

[2] If the price $p_t \neq \tilde{v}$ for some $t < T$, and the agent did not trade in $[t, T)$, there would have to be a jump in the price at time T, which the results of our model rule out. This would not be rational for the insider to do, as she would miss some profit opportunities by not trading.

The corresponding price p_t set by the market makers is

$$p_t = E[\tilde{v}|\mathcal{F}_t^{\hat{y}}] = \frac{p_0 + S_0 \int_0^t \frac{\beta_s}{\sigma_s^2} d\hat{y}_s}{1 + S_0 \int_0^t (\frac{\beta_s}{\sigma_s})^2 ds} \tag{2.12}$$

$$= E[\tilde{v}] + \int_0^t \lambda_s dy_s,$$

where the price sensitivity λ_t is given by

$$\lambda_t = \Big[\frac{S_0}{\int_0^T \sigma_s^2 ds}\Big]^{1/2}. \tag{2.13}$$

The corresponding mean square error is

$$S_t(\beta) := E[(\tilde{v} - p_t)^2] = \frac{S_0 \int_t^T \sigma_s^2 ds}{\int_0^T \sigma_s^2 ds}; \qquad t \in [0, T]. \tag{2.14}$$

In particular, $S_T(\beta) = 0$, which by (2.9) implies that

$$\tilde{v} = p_T \quad a.s. \tag{2.15}$$

3 Properties of the equilibrium.

The generalization relative to Kyle (1985) included in Theorem 2.1 allows for a time varying volatility parameter in the order process of the noise traders. One would, perhaps, expect that as a consequence the market liquidity function λ_t would depend on time, suggested by the expression (4.39) in the next section. The result of Theorem 2.1 is that it does not. The intuition for this can be explained as follows:

The trading intensity β_t will typically increase as t approaches T, since the insider becomes increasingly desperate to utilize his residual information advantage. In particular, from expression (2.10) in Theorem 2.1 we see that β_t/σ_t^2 increases as t increases. It follows from the proof in the next section, equations (4.38) and (4.39), that the price sensitivity λ_t can be written

$$\lambda_t = \frac{\beta_t S_t}{\sigma_t^2}.$$

By the well-known Kalman-Bucy filter we have

$$\frac{dS_t}{dt} = -\Big(\frac{\beta_t}{\sigma_t} S_t\Big)^2, \qquad \text{where } S_t = S_t(\beta). \tag{3.1}$$

Solving this equation we see that S_t has the form

$$S_t = \frac{S_0}{1 + S_0 \int_0^t \tilde{\beta}_s^2 ds}; \qquad t \in [0, T],$$

where

$$\tilde{\beta}_t = \frac{\beta_t}{\sigma_t}; \qquad 0 \leq t \leq T.$$

The quantity $\int_0^t \tilde{\beta}_s^2 ds$ measures the the "amount" of insider trading to liquidity trading by time t. As this quantity increases over time, the amount of private information S_t remaining at time t is seen, from the above expression, to decrease, where S_t is the (mean square) distance between $\tilde{v}$ and p_t. It follows from the proof in Section 4 that if β is optimal, then (see (4.35))

$$S_t = \frac{S_0 \int_t^T \sigma_s^2 ds}{\int_0^T \sigma_s^2 ds}.$$

From this we conclude that if β is optimal, then not only does S_t decrease over time, meaning that the insider's information gradually enters the price p_t, but also

$$S_T = 0 \qquad \text{and hence } \; p_T = \tilde{v} \text{ a.s.}$$

The function λ_t is seen to depend on two effects:
(i) The quantity β_t/σ_t^2 increases over time, which tends to increase λ_t as time t increases.
(ii) The quantity S_t decreases over time, suggesting that the insider's information advantage is deteriorating, which tends to decrease λ_t as t increases.
In equilibrium (i) is offset by (ii) and $\lambda_t = \lambda$ is constant over time.

Notice that the important quantities are β_t/σ_t^2 and $\beta_t/\sigma_t = \tilde{\beta}_t$ in the above arguments. The mere fact that the amount of insider trading represented by $\int_0^t \beta_s^2 ds$ is large, is no guarantee that the market price p_t is close to the fundamental value $\tilde{v}$, i.e., that S_t is small. It could be that the amount of noise trading $\int_0^t \sigma_s ds$ is also large, in which case the insider could hide his trade, and less information about the true value would be revealed to the market makers. Similarly, we do not know that β_t is *monotonically* increasing over time, only that β_t/σ_t^2 is. Notice that the equilibrium value of the price sensitivity λ can be interpreted as the square root of a ratio, where the numerator is the amount of private information, ex ante, and the denominator is the amount of liquidity trading.

From the expressions in Theorem 2.1 we notice that

$$\beta_t = \frac{1}{\lambda} \frac{\sigma_t^2}{\int_t^T \sigma_s^2 ds}$$

so β_t is inversley related to λ for each t. Since the quantity $1/\lambda$ measures the market depth, the insider will naturally trade more intensely, ceteris paribus, when this quantity is large.

From the general discussion in Kyle (1985) it is indicated that if the slope of the residual supply curve λ_t ever decreases (i.e., if the market depth ever increases), then unbounded profits can be generated. This is inconsistent with an equilibrium, so λ_t must be monotonically non-decreasing in any equilibrium. It is argued that this follows since in continuous time, the informed trader can act as a perfectly discriminating monopsonist, moving up or down the residual supply curve (i.e., the market is infinitely tight). Hence, she could exploit predictable shifts in the supply curve. From the analysis of Back (1992) it is known that, more generally, this slope must be a martingale given the market makers' information. Our result that λ_t is indeed a constant is, accordingly, consistent with the literature.

One would, perhaps, expect that the insider, since she knows the function σ_t, may use it to further conceal her trade in that she will use a high β_t at a time when σ_t is large. This impression is confirmed by investigating the optimal trading intensity β appearing in expression (2.10) of Theorem 2.1.

However, when σ_t is low the insider must apply a correspondingly lower trading intensity, and it turns out that the expected (ex ante) profits average out. This can be demonstrated as follows: Consider the expected wealth of the insider

$$E[w_T] = w_0 + S_0 \int_0^T \frac{\beta_t dt}{1 + S_0 \int_0^t \tilde{\beta}_s^2 ds},$$

an expression which follows from the results of the next section. Here the last term is the expected (ex ante) profits, which can be shown to be

$\sqrt{S_0 \int_0^T \sigma_t^2 dt}$.[3] Thus, trading at a time-varying volatility σ_t corresponds exactly, when it comes to expected profits, to trading at a constant volatility σ determined by $\sigma^2 = \frac{1}{T}\int_0^T \sigma_t^2 dt$, the right comparison in this regard.

The explanation is that in this model both the insider and the market makers can be assumed to know the value of σ_t at any time t. Accordingly the insider cannot utilize the variability in this volatility to further conceal her trades, and thus make additional profits

When the amount of liquidity trading $\int_0^t \sigma_s^2 ds$ is large, we noticed above that λ is small, in which case the insider's profit is large. However, a small value of λ is, in isolation, no guarantee for a large ex ante profit of the insider, since a large value of S_0 also makes the profit of the insider large, and λ large as well.

This points in one possible direction for extending the present model. Suppose that the private information is connected to quaterly accounting data for the firm, so T stands for one quarter, and let us extend the model beyond T to $2T$, $3T, \cdots$, etc. Let us, as in Admati and Pfleiderer (1988), imagine two types of liquidity traders, discretionary and non-discretionary. Just after each disclosure period of length T, the level of private information relative to the uninformed is at its minimum. It seems reasonable, from the above formula for the ex ante profits of the insider, that the discretionary traders, acting strategically to time their trades, should concentrate their trade to these times in order to loose less to the insider. That this kind behavior is optimal is expected from the conclusions of Admati and Pfleiderer (1988), who noticed that λ is a constant is not in accordance with empirical findings; the bid ask spread 2λ is varying over time.

We also have the following corollary:

Corollary 3.1. *Suppose $\sigma_t = \sigma > 0$ is a constant. Then the optimal trading intensity for the insider is*

$$\beta_t = \frac{\sigma\sqrt{T}}{\sqrt{S_0}(T-t)}; \qquad 0 \le t < T. \tag{3.2}$$

The corresponding price p_t set by the market makers is given by

$$dp_t = \lambda_t dy_t, \tag{3.3}$$

[3] In the case when $\sigma_t = \sigma$ is a constant, we get that the expected profits equal $\sigma\sqrt{S_0 T}$, consistent with Kyle (1985).

where

$$\lambda_t \equiv \lambda = \frac{\sqrt{S_0}}{\sigma}\frac{1}{\sqrt{T}}; \qquad \text{a constant for all } t \in [0,T). \tag{3.4}$$

This result follows from Theorem 2.1 by setting $\sigma_s \equiv \sigma$ in (**??**). The results of Corollary 3.1 are in agreement with Kyle (1985) and Back (1992) (when we set $T = 1$).

Recently, a paper of related interest by Eide (2007) came to our knowledge. Her work, which was done independently of ours, differs from ours in several ways: She focuses on the situation when the price process $\tilde{v}_t$ of the stock is assumed to have a specific dynamics (an Itô diffusion and a martingale with respect to an independent Brownian motion), and its current value $\tilde{v}_t$ (not $\tilde{v}_T$) is known to the insider at time t for all $t \in [0,T]$. She avoids the use of forward integrals by assuming a priori that the processes are semimartingales with respect to the relevant filtrations. Like Back she then assumes that the market makers set the price equal to $p_t = H(t, y_t)$ for some function H and that $H(t, y_t) = E(\tilde{v}_T|\mathcal{F}_t^y)$. These assumptions put the problem of finding a corresponding equilibrium into a Markovian context, which allows her to solve the problem by using dynamic programming. In conclusion, her a priori assumptions are stronger than ours, but they enable her to solve other problems than we do. In particular, the final stock value $\tilde{v} = \tilde{v}_T$ need not be normally distributed in her case.

Remark 3.2. To summarize, our paper differes from the papers of Kyle (1985) and Back (1992) both with respect to basic assumptions and method:

(i) We do *not* assume that the volatility $\sigma(t)$ of the noise traders is constant. Nevertheless we prove that the price sensitivity λ_t is constant also in our case, if the optimal strategy is applied.

(ii) We do *not* assume a priori that

$$p_T = \tilde{v} \quad \text{a.s.}$$

But this is *proved* to be the case if the optimal strategy is used.

We remark that if we had made this assumption a priori, then our proof could have been simplified as follows: The last term in (4.14) would have been 0. Hence (see (4.16)) we would have $S_{t,T}^{(\beta)} = 0$ for all $t \in [0,T]$ and Problem 1 would automatically reduce to Problem 2.

(iii) We do *not* assume a priori that the strategy x_t is *incospicuous*, i.e. that

$$\frac{1}{\sigma_t} dy_t = \frac{1}{\sigma_t} x_t dt + dz_t$$

is a Brownian motion with respect to its own filtration. However, this is *proved* to hold if x_t is chosen optimally. [4]

(iv) We do *not* assume a priori that there exists a function H such that

$$p_t = H(t, y_t).$$

But this is *proved* to be the case if the insider acts optimally.

(v) Finally, since we are not assuming a Markovian setup we cannot use dynamic programming (the HJB equation) to find the optimal strategy, but we use filtering theory and a perturbation argument instead.

Remark 3.3. It is interesting to note that also in our general setting the total order process y_t becomes a *Brownian bridge* with respect to the filtration $\mathcal{G}_t$ if the optimal insider strategy is used. To see this we proceed as follows:

By (2.7), (2.8), (2.9) we have

$$\begin{aligned} dy_t &= (\tilde{v} - p_t)\beta_t dt + \sigma_t dB_t \\ &= (\tilde{v} - E[\tilde{v}] - \lambda y_t)\beta_t dt + \sigma_t dB_t \\ &= \Big[\Big(\frac{\int_0^T \sigma_u^2 du}{S_0}\Big)^{1/2} (\tilde{v} - E[\tilde{v}]) - y_t\Big] \frac{\sigma_t^2 dt}{\int_t^T \sigma_u^2 du} + \sigma_t dB_t \,. \end{aligned} \tag{3.5}$$

Thus y_t is the *bridge* of the process $z_t = \int_0^t \sigma_s dB_s$, conditioned to arrive at the terminal value

$$y_T = \Big(\frac{\int_0^T \sigma_u^2 du}{S_0}\Big)^{1/2} (\tilde{v} - E[\tilde{v}])$$

at time $t = T$.

In particular, if $\sigma_t = \sigma$ is constant we get

$$dy_t = \Big[\sigma\Big(\frac{T}{S_0}\Big)^{1/2} (\tilde{v} - E[\tilde{v}]) - y_t\Big] \frac{dt}{T - t} + \sigma \, dB_t, \tag{3.6}$$

and hence $\frac{1}{\sigma} dy_t$ is the classical Brownian bridge, conditioned to arrive at

$$\Big(\frac{T}{S_0}\Big)^{1/2} (\tilde{v} - E[\tilde{v}])$$

at time $t = T$.

[4] Also Back (1992) shows this, using a different method.

4 The solution of the problem

In this section we present the proof of Theorem 2.1. It can be noted to be rather different from the corresponding development in Kyle (1985).

To summarize the model mathematically, the portfolio of the noise traders has the form

$$dz_t = \sigma_t dB_t\,, \quad t \in [0,T], \tag{4.1}$$

and the portfolio of the insider is

$$dx_t = (\tilde{v} - p_t)\beta_t dt\,, \tag{4.2}$$

where p_t is the market price at time t set by the market makers. The total traded volume is hence

$$dy_t = (\tilde{v} - p_t)\beta_t dt + \sigma_t dB_t\,. \tag{4.3}$$

If we let $\mathcal{F}_t^y$, $t \in [0,T]$, be the filtration generated by y_s; $s \le t$, then it is assumed that

$$p_t := E[\tilde{v}|\mathcal{F}_t^y], \quad 0 \le t \le T. \tag{4.4}$$

Substituting this into (4.3) we get that the total traded volume process must satisfy the equation

$$dy_t = (\tilde{v} - E[\tilde{v}|\mathcal{F}_t^y])\beta_t dt + \sigma_t dB_t\,, \quad t \in [0,T]\,. \tag{4.5}$$

Thus, it is an assumption of the whole setup that a solution process y_t of this (highly non-standard) equation (4.5) exists. The main idea of our approach is that we prove that it is possible to find a solution of (4.5) by regarding y_t as the *innovation process* $\tilde{y}_t$ of an auxiliary linear filtering problem, where the *signal process* is

$$d\tilde{v}_t = 0, \tilde{v}_0 = \tilde{v}; \quad t \in [0,T], \tag{4.6}$$

and the *observation process* is

$$d\hat{y}_t = \tilde{v}\beta_t dt + \sigma_t dB_t\,; \quad t \in [0,T],\ \hat{y}_0 = 0\,. \tag{4.7}$$

The innovation process for this problem is, by definition,

$$\begin{aligned} d\tilde{y}_t &= (\tilde{v} - E[\tilde{v}|\mathcal{F}_t^{\hat{y}}])\beta_t dt + \sigma_t dB_t \\ &= d\hat{y}_t - E[\tilde{v}|\mathcal{F}_t^{\hat{y}}]\beta_t dt\,, \end{aligned} \tag{4.8}$$

where $\mathcal{F}_t^{\hat{y}} = \sigma(\hat{y}_s\,,\ 0 \le s \le t)$ is the information filtration generated by $\hat{y}$.

As before let $\mathcal{F}_t^y = \sigma(y_s; s \le t)$ be the information filtration of the process y. Then we have:

Lemma 4.1. $\mathcal{F}_t^y = \mathcal{F}_t^{\hat{y}}$*;* $\quad t \in [0,T]$.

Proof. The proof of Lemma 6.2.5 (iii) in Øksendal (2003) applies without changes. □

Corollary 4.2. *The innovation process $\tilde{y}_t$ is a solution of the equation* (4.5) *for the total traded volume process y_t.*

Based on this we choose the innovation process $\tilde{y}_t$ to represent the total order process y_t and we write $\tilde{y}_t = y_t$ from now on.

Note that from filtering theory we know that the process y^* defined by $dy_t^* := \frac{1}{\sigma_t} dy_t$ is a Brownian motion with respect to the information filtration $\mathcal{F}_t^y$. [5]

As before let

$$S_t = S_t^{(\beta)} := E[(\tilde{v} - p_t)^2] \tag{4.9}$$

be the mean square error process and define

$$S_{t,T} = S_{t,T}^{(\beta)} := E[(\tilde{v} - p_t)(\tilde{v} - p_T)]; \qquad 0 \le t \le T. \tag{4.10}$$

(Note that if we had assumed that

$$p_T = \tilde{v} \quad \text{a.s.}$$

[5] Back (1992) also has this result using a different method.

then we would get $S_{t,T} = 0$ and the following proof would simplify considerably.)

Then (2.7) can be written

$$E[w_T] = w_0 + \int_0^T S_t^{(\beta)} \beta_t dt - \int_0^T S_{t,T}^{(\beta)} \beta_t dt. \tag{4.11}$$

We need to compute $S_{t,T}^{(\beta)} = E[(\tilde{v} - p_T)(\tilde{v} - p_t)]$: We have

$$\begin{aligned} E[(\tilde{v} - p_T)(\tilde{v} - p_t)] &= E[(\tilde{v}^2) - E[(\tilde{v}p_t) - E(\tilde{v}p_T) + E(p_T p_t) \\ &= E(\tilde{v}^2) - E(p_t^2) - E(p_T^2) + E(p_T p_t)\,. \end{aligned}$$

We first compute $E(p_T p_t)$. By (4.4) we have that p_t is a square-integrable martingale. Hence

$$E[p_t p_T] \;=\; E[p_t^2],$$

and consequently

$$\begin{aligned} E[(\tilde{v} - p_T)(\tilde{v} - p_t)] &= E(\tilde{v}^2) - E(p_t^2) - E(p_T^2) + E(p_T p_t) \\ &= E(\tilde{v}^2) - E(p_t^2) - E(p_T^2) + E(p_t^2) \\ &= E(\tilde{v}^2) - E(p_T^2)\,. \end{aligned}$$

But

$$E(p_T^2) = E(\tilde{v}^2) - E(\tilde{v} - p_T)^2 = E(\tilde{v}^2) - S(T)\,,$$

and hence

$$S_{t,T}^{(\beta)} \;= E[(\tilde{v} - p_T)(\tilde{v} - p_t)] \;= S_T(\beta)\,. \tag{4.12}$$

In particular, note that

$$S_{t,T}^{(\beta)} \geq 0 \qquad \text{for all } t \in [0, T] \tag{4.13}$$

and

$$S_{t,T}^{(\beta)} = 0 \qquad \text{if } p_T = \tilde{v}. \tag{4.14}$$

We now return to problem (2.8). By (3.17) we see that our original problem can be formulated as the following control problem:

Problem 4.3. Maximize

$$J_1(\beta) := \int_0^T S_t(\beta)\beta_t dt - S_T(\beta)\int_0^T \beta_t dt \tag{4.15}$$

over all $\beta \in \mathcal{A}$, where $\mathcal{A}$ is the set of all (deterministic) functions $\beta : [0,T) \to \mathbb{R}$ which are continuous on $[0,T)$.

We first study the following related problem:

Problem 4.4. Maximize

$$J(\beta) := \int_0^T S_t(\beta)\beta_t dt \tag{4.16}$$

over all $\beta \in \mathcal{A}$.

We will find the optimal control $\hat{\beta} \in \mathcal{A}$ for Problem 4.3 and show that the corresponding terminal price $p_T^{(\hat{\beta})}$ satisfies

$$p_T^{(\hat{\beta})} = \tilde{v} \quad \text{a.s.} \tag{4.17}$$

It follows by (4.15) that $S_{t,T}^{(\hat{\beta})} = S_T(\hat{\beta}) = 0$ and hence $\hat{\beta}$ is also optimal for Problem 4.3, because,

$$\sup_{\beta \in \mathcal{A}} J_1(\beta) \leq \sup_{\beta \in \mathcal{A}} J_1(\beta) = J(\hat{\beta}) = J_1(\hat{\beta}) \leq \sup_{\beta \in \mathcal{A}} J_1(\beta).$$

The first inequality holds since $J_1(\beta) \leq J(\beta)$ for all β. (We assume that $\beta \neq 0$.) The second (in)equality holds by the definition of $\hat{\beta}$. The third (in)equality holds since $S_{t,T}^{(\hat{\beta})} = 0$. The fourth inequality holds since $\hat{\beta}$ is just one of possible β's in the maximum.

In view of this we now proceed to solve Problem 4.3. Since the map

$$\beta \to J(\beta); \qquad \beta \in \mathcal{A}$$

is concave, we can use the following perturbation argument to find the maximizer for $J(\cdot)$:

Suppose $\beta \in \mathcal{A}$ maximizes $J(\beta)$. Choose an arbitrary function $\xi \in \mathcal{A}$ and define the real function g by

$$g(y) = J(\beta + y\xi), \qquad y \in \mathbb{R}.$$

Then g is maximal at $y = 0$ and hence

$$
\begin{aligned}
0 = g'(0) &= \frac{d}{dy} J(\beta + y\xi)|_{y=0} \\
&= \frac{d}{dy}\Big(\int_0^T S_t(\beta + y\xi)(\beta_t + y\xi_t) dt \Big)\Big|_{y=0} \\
&= I_1 + I_2, \qquad (4.18)
\end{aligned}
$$

where

$$I_1 = \int_0^T S_t(\beta)\xi_t dt \tag{4.19}$$

and

$$I_2 = \int_0^T \beta_t \frac{d}{dy} S_t(\beta + y\xi)|_{y=0} dt. \tag{4.20}$$

Define

$$\eta_t = \frac{d}{dy} S_t(\beta + y\xi)|_{y=0}. \tag{4.21}$$

By the well-known Kalman-Bucy filter we have

$$\frac{dS_t}{dt} = -\Big(\frac{\beta_t}{\sigma_t} S_t\Big)^2, \qquad \text{where } S_t = S_t(\beta). \tag{4.22}$$

Hence

$$S_t = S_0 - \int_0^t \Big(\frac{\beta_s}{\sigma_s} S_s\Big)^2 ds.$$

Therefore

$$
\begin{aligned}
\eta_t &= -\int_0^t \frac{d}{dy}\Big[\Big(\frac{\beta_s + y\xi_s}{\sigma_s} S_s(\beta + y\xi)\Big)^2\Big]_{y=0} ds \\
&= -\int_0^t 2\Big(\frac{\beta_s}{\sigma_s} S_s(\beta)\Big)\Big[\frac{\xi_s}{\sigma_s} S_s(\beta) + \frac{\beta_s}{\sigma_s}\eta_s\Big] ds.
\end{aligned}
$$

Differentiating with respect to t we get

$$\frac{d\eta_t}{dt} = -\frac{\gamma_t \xi_t}{\sigma_t} S_t(\beta) - \frac{\gamma_t \beta_t}{\sigma_t}\eta_t$$

where

$$\gamma_t = 2\frac{\beta_t}{\sigma_t} S_t(\beta). \tag{4.23}$$

Hence

$$\frac{d\eta_t}{dt} + \frac{\gamma_t \beta_t}{\sigma_t}\eta_t = -\frac{\gamma_t \xi_t}{\sigma_t} S_t(\beta).$$

Multiplying by $\exp\big(\int_0^t \frac{\gamma_r\beta_r}{\sigma_r}dr\big)$ we obtain

$$\frac{d}{dt}\Big(\eta_t \exp\Big(\int_0^t \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)\Big) = -\frac{\gamma_t\xi_t}{\sigma_t} S_t(\beta)\exp\Big(\int_0^t \frac{\gamma_r\beta_r}{\sigma_r}dr\Big).$$

Note that

$$\eta_0 = \frac{d}{dy} S_0(\beta + y\xi)|_{y=0} = \frac{d}{dy} E[(\tilde{v} - E[\tilde{v}])^2] = 0.$$

Hence, by integrating the above,

$$\eta_t = -\exp\Big(-\int_0^t \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)\int_0^t \frac{\gamma_s\xi_s}{\sigma_s} S_s(\beta)\exp\Big(\int_0^s \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)ds. \tag{4.24}$$

Substituting this in (4.20) and changing the order of integration we get

$$\begin{aligned} I_2 &= \int_0^T \beta_t\eta_t dt \\ &= -\int_0^T \beta_t\Big[\int_0^t \frac{\gamma_s\xi_s}{\sigma_s} S_s(\beta)\exp\Big(-\int_s^t \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)ds\Big]dt \\ &= -\int_0^T \Big[\int_s^T \beta_t \exp\Big(-\int_s^t \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)dt\Big]\frac{\gamma_s\xi_s}{\sigma_s} S_s(\beta)ds. \end{aligned}$$

Changing the notation between s and t we get

$$I_2 = -\int_0^T \Big[\int_t^T \beta_s \exp\Big(-\int_t^s \frac{\gamma_r\beta_r}{\sigma_r}dt\Big)ds\Big]\frac{\gamma_t S_t(\beta)}{\sigma_t}\xi_t dt. \tag{4.25}$$

Combining this with (4.18) and (4.19) we obtain

$$\int_0^T \Big\{S_t(\beta) - \Big[\int_t^T \beta_s \exp\Big(-\int_t^s \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)ds\Big]\frac{\gamma_t}{\sigma_t} S_t(\beta)\Big\}\xi_t dt = 0.$$

Since this holds for all $\xi \in \mathcal{A}$ we conclude that

$$S_t(\beta) - \Big[\int_t^T \beta_s \exp\Big(-\int_t^s \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)ds\Big]\frac{\gamma_t}{\sigma_t}S_t(\beta) = 0; \qquad t \in [0,T]. \tag{4.26}$$

Recall that we have assumed that (see (2.9))

$$S_t(\beta) > 0 \qquad \text{for all } t \in [0,T). \tag{4.27}$$

Hence (4.26) implies that

$$\Big[\int_t^T \beta_s \exp\Big(-\int_t^s \frac{\gamma_r\beta_r}{\sigma_r}dr\Big)ds\Big]\frac{\gamma_t}{\sigma_t} = 1; \qquad t \in [0,T). \tag{4.28}$$

From this we deduce that

$$\lim_{s\to T^-} \beta_s = \infty \quad \text{or} \quad \lim_{t\to T^-} \frac{\gamma_t}{\sigma_t} = \infty\,, \quad \text{or both.} \tag{4.29}$$

By (4.28) we see that in either case we can deduce that

$$\lim_{t\to T^-} \beta_t = \infty\,. \tag{4.30}$$

Put

$$u(t) = \frac{\gamma_t\beta_t}{\sigma_t}\,, \qquad v(t) = \int_0^t u(r)dr. \tag{4.31}$$

Then (4.28) gives

$$\int_t^T \beta_s \exp(-v(s))ds = \frac{\beta_t}{u(t)}\exp(-v(t)).$$

Differentiating we get

$$-\beta_t \exp(-v(t)) = \Big[\frac{d}{dt}\Big(\frac{\beta_t}{u(t)}\Big) - \frac{\beta_t u(t)}{u(t)}\Big]\exp(-v(t))$$

or

$$\frac{d}{dt}\Big(\frac{\beta_t}{u(t)}\Big) = 0; \qquad t \in [0,T).$$

From this we deduce that

$$u(t) = C_1\beta_t; \qquad t \in [0,T)$$

i.e.

$$\gamma_t = C_1\sigma_t; \qquad t \in [0,T)$$

for some constant C_1. Hence, by (4.23)

$$\frac{\beta_t}{\sigma_t} S_t(\beta) = C_2\sigma_t, \qquad t \in [0,T) \tag{4.32}$$

where $C_2 = \frac{1}{2}C_1$.

We conclude that the optimal β_t must satisfy the equation

$$\beta_t = \frac{C_2\sigma_t^2}{S_t(\beta)}. \tag{4.33}$$

Hence, by (4.30)

$$S_T(\beta) = \lim_{t\to T^-} S_t(\beta) = 0. \tag{4.34}$$

Moreover, by (4.22) and (4.32),

$$\begin{aligned} \frac{d}{dt} S_t(\beta) &= -\Big(\frac{\beta_t}{\sigma_t} S_t(\beta)\Big)^2 \\ &= -C_2^2\sigma_t^2, \end{aligned}$$

which integrates to

$$S_t(\beta) = S_T(\beta) + C_2^2 \int_t^T \sigma_s^2 ds = C_2^2 \int_t^T \sigma_s^2 ds.$$

Choosing $t = 0$ we get

$$C_2 = \Big[\frac{S_0}{\int_0^T \sigma_s^2 ds}\Big]^{1/2}.$$

Hence, $\beta = \beta^*$ is optimal iff

$$S_t(\beta) = \frac{S_0 \int_t^T \sigma_s^2 ds}{\int_0^T \sigma_s^2 ds} \tag{4.35}$$

and the optimal $\beta = \beta^*$ is given explicitly by

$$\beta_t = \frac{S_0^{1/2}(\int_0^T \sigma_s^2 ds)^{1/2}\sigma_t^2}{S_0 \int_t^T \sigma_s^2 ds}; \qquad t \in [0,T). \tag{4.36}$$

This gives that the maximal value $J(\beta^*)$ of $J(\beta)$ is

$$J(\beta) = \int_0^T S_t(\beta)\beta_t dt$$

$$= \left[S_0 \int_0^T \sigma_s^2 ds\right]^{1/2} \tag{4.37}$$

and hence that the maximal expected terminal wealth of the insider is

$$E[w_T] = w_0 + \left[S_0 \int_0^T \sigma_s^2 ds\right]^{1/2}. \tag{4.38}$$

Finally, by the Kalman-Bucy filter the corresponding filtered estimate p_t is given by

$$p_t = E[\tilde{v}] + \int_0^t \lambda_s dy_s; \qquad t \in [0,T], \tag{4.39}$$

where the price sensitivity λ_t is given by

$$\lambda_t = \frac{S_t(\beta)\beta_t}{\sigma_t^2} = \left[\frac{S_0}{\int_0^T \sigma_s^2 ds}\right]^{1/2}; \qquad t \in [0,T]. \tag{4.40}$$

This concludes the proof of Theorem 2.1.

Acknowledgments We want to thank Francesca Biagini, Albina Danilova, Yaozhong Hu, Kjell Henry Knivsflå, Thilo Meyer-Brandis and Dirk Paulsen for valuable comments.

References

[1] Aase,K., Bjuland,T. and Øksendal,B. (2010). "An anticipative linear filtering equation". Manuscript 2010.

[2] Admati,A.R. and Pfleiderer,P. (1988). "A Theory of Intraday Patterns: Volume and Price Variability". *The Review of Financial Studies* 1, 1, 3–40.

[3] Allinger,D.F. and Mitter,S.K. (1981). "New Results on the Innovations Problem for Non-Linear Filtering". *Stochastics* 4, 339–348.

[4] Back,K. (1992). "Insider Trading in Continuous Time". *The Review of Financial Studies* Vol. 5, No 3, 387-409.

[5] Biagini,F., and Øksendal,B. (2005). "A general stochastic calculus approach to insider trading". *Appl. Math. Optim.* 52, 167–181.

[6] Davis,M.H.A. (1977). *Linear Estimation and Stochastic Control.* Chapman and Hall.

[7] Davis,M.H.A. (1984). *Lectures on Stochastic Control and Nonlinear Filtering.*. Tata Institute of Fundamental Research, Bombay.

[8] Eide,I.B. (2007). *"An equilibrium model for gradually revealed asymmetric information"*. Preprint, University of Oslo 6/2007.

[9] Glosten,L.R., and Milgrom,P.R. (1985). "Bid, Ask and Transaction Prices in a Specialist Market with Hetrogeneously Informed Traders". *Journal of Financial Economics*, 14, 71–100.

[10] Grossman,S.J., and Stiglitz,J.E. (1980). "On the Impossibility of Informationally Efficient Markets". *American Economic Review*, 70 , 393–408.

[11] Holden,C.W. and Subrahmanyam,A. (1992). "Long-Lived Private Information and Imperfect Competition". *The Journal of Finance*, XLVII, 1 247–270.

[12] Kallianpur,G. (1980). *Stochastic Filtering Theory.* Springer.

[13] Kalman,R.E. (1960). "A new approach to linear filtering and prediction problems". *J. Basic Engineering* D 82, 35–45

[14] Kyle,A.S. (1985). "Continuous Auctions and Insider Trading". *Econometrica* Vol.53, No. 6, 1315–1336.

[15] Liptser,R.S. and Shiryaev,A.N.: Statistics of Random Processes II. Springer 1978.

[16] Øksendal,B. (2003). Stochastic Differential Equations. 6th Edition. Springer

[17] Russo,F., and Vallois,P. (1993). "Forward, backward and symmetric stochastic integration". *Probab. Theory Related Fields* 97, 403–421.

[18] Russo,F. and Vallois,P. (1995). "The generalized covariation process and Itô formula". *Stoch. Process. Appl.* 59, 81–104.

[19] Russo,F. and Vallois,P. (2000). "Stochastic calculus with respect to continuous finite quadratic variation processes". *Stoch. Stoch. Rep.* 70, 1–40.