PARTIAL AND ASYMMETRIC INFORMATION

KERRY BACK
JOHN M. OLIN SCHOOL OF BUSINESS
WASHINGTON UNIVERSITY IN ST. LOUIS
ST. LOUIS, MO 63130
BACK@OLIN.WUSTL.EDU

This is a preliminary and incomplete version of a manuscript to be prepared for the CIME-EMS Summer School on Stochastic Methods in Finance, July, 2003. Please do not quote or distribute.

1. Introduction

The term "partial information" could mean one of two things: either "incomplete information" or "advance information." The study of incomplete information is an application of filtering theory. The study of advance information, which may be considered a study of "insider trading," is an application of the theory of enlargement of filtrations. In one sense, these two topics are the opposite of each other: in the study of incomplete information one starts with the filtration on which all processes are adapted and assumes investors have access to only a subfiltration, and in the study of advance information one starts with a process on which the return processes of assets are adapted and assumes some investor or investors have access to a larger filtration. However, in another sense, the two topics are the same: in both cases, the change of filtration simply changes the drifts of economic processes, e.g., the expected returns of assets.

The second part of the title, "asymmetric information," means that some investors have information in advance of others. The study of asymmetric information, which emphasizes differences in information, means that we will be concerned with equilibrium theory and how the less informed agents learn in equilibrium from the more informed agents. The study of incomplete information is also most interesting in the context of economic equilibrium. In these lectures, I will focus primarily on incomplete information and asymmetric information, treating the topic of advance information within the context of asymmetric information (see Section 4.1).

Problems of control under incomplete information are characterized by the separation principle: one first estimates the unobserved process and then chooses the optimal control. Portfolio choice problems under incomplete information satisfy this principle. Problems of asymmetric information are necessarily also problems of incomplete information, since "asymmetry" means that some agents do not possess the information of others. Hence, what is unobserved by one agent may be observed by another. Financial equilibrium occurs when the control problems solved by agents have consistent solutions, in the sense that the optimal controls satisfy market-clearing conditions. Because the controls of one agent (or one class of agents) may affect what is observed by other agents, the filtering and control problems of various agents may be intertwined. This means that, though each individual control problem satisfies the separation principle, equilibrium can be computed only by solving the various filtering and control problems simultaneously.

Excellent surveys of incomplete information models in finance and of asymmetric information models have recently been published. Ziegler (2003) surveys incomplete information models, and Brunnermeir (2001) surveys asymmetric information models. In these lectures, I will not attempt to repeat these comprehensive surveys but instead will give a more selective review.

2. Filtering Theory

Let us start with a brief review of filtering theory, developed in the 1970's, and exposited, for example, in Rogers and Williams (2000). Note first that engineers and economists use the term "signal" differently. Engineers take the viewpoint of the transmitter, who sends the "signal," which is then to be estimated (or "filtered") from a noisy observation.. Economists generally take the viewpoint of the receiver, who observes a "signal" about some other variable and then uses the signal to estimate the latter variable. To avoid confusion, I will try to avoid the term, but when I slip and use it, it will be in the sense that economists use it.

All processes are adapted to some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$. The problem is to estimate a process X from the observations of another process Y. Specifically, we consider estimating the conditional expectation $E[f(X_t)|\mathcal{F}_t^Y]$, where $\{\mathcal{F}_t^Y\}$ is the usual augmentation of the filtration generated by Y, and f is a real-valued function satisfying some minimal regularity conditions but otherwise arbitrary. By estimating $E[f(X_t)|\mathcal{F}_t^Y]$ for arbitrary f, one can obtain the distribution of X_t conditional on $\mathcal{F}_t$.

Assume

$$dY_t = h_t\,dt + dW_t; \quad Y_0 = 0 \tag{2.1}$$

where W is a Brownian motion on $\mathbb{R}^n$ and h is an $\mathbb{R}^n$–valued process. Define $f_t = f(X_t)$ and assume

$$df_t = g_t\,dt + dM_t, \tag{2.2}$$

for some process g, where M is a martingale. We follow the convention of denoting the $\{\mathcal{F}_t^Y\}$–optional projection of a process with the "hat" symbol. So we want to compute $\hat{f}_t$.

The "innovation process" is defined as

$$\begin{aligned} dZ_t &= dY_t - \hat{h}_t\,dt \\ &= (h_t - \hat{h}_t)\,dt + dW_t \end{aligned} \tag{2.3}$$

with $Z_0 = 0$. The differential dZ is interpreted as the innovation or "surprise" in the variable Y, which consists of two parts, one being the error in the estimation of the drift h_t and the other being the random change dW.

The main results of filtering theory, due to Fujisaki, Kallianpur and Kunita (1972), are the following:

1) The innovation process Z is an $\{\mathcal{F}_t^Y\}$–Brownian Motion on $\mathbb{R}^n$.
2) For any L^2–bounded $\{\mathcal{F}_t^Y\}$–martingale H, there exist $\{\mathcal{F}_t^Y\}$–previsible processes ϕ^i such that $E\int_0^T (\phi_t^i)^2\,dt < \infty$ for $i = 1,\ldots,n$, and
$$dH_t = \sum_{i=1}^{n} \phi_t^i\,dZ_t^i.$$
3) There exist processes α^i such that $d[M,W^i]_t = \alpha_t^i\,dt$, for $i = 1,\ldots,N$.
4) $\hat{f}$ evolves as
$$d\hat{f}_t = \hat{g}_t\,dt + \sum_{i=1}^{N}\left(\widehat{fh}_t - \hat{f}_t\hat{h}_t + \hat{\alpha}_t\right)\,dZ_t^i,$$
where $\widehat{fh}_t$ denotes the optional projection of the process fh.

Part (1) means in particular that Z is a martingale; thus the innovations dZ are indeed "unpredictable." Given that it is a martingale, the fact that it is a Brownian motion follows from Levy's theorem and the fact, which follows immediately from (2.3), that the covariations are $d\langle Z^i, Z^j\rangle = dt$ if $i = j$ and 0 otherwise. Part (2) means that the process Z "spans" the $\{\mathcal{F}_t^Y\}$–martingales (which would follow from $\{\mathcal{F}_t^Y\}$ $= \{\mathcal{F}_t^Z\}$, though that need not be true in general). Part (3) means that the covariation processes are absolutely continuous, though in our applications we will assume M and the W^i are independent, implying $\alpha^i = 0$ for all i.

Part (4) is the filtering formula. The estimate $\hat{f}$ is updated because f is expected to change (which is obviously captured by the term $\hat{g}_t\,dt$) and because new information from dZ is available to estimate f. The observation process Y (or equivalently the innovation process Z) is useful for estimating f due to two factors. One is the possibility of correlation between the martingales W and M. This reflected in the term $\hat{\alpha}_t\,dZ_t$. The other factor is the correlation between f and the drift h_t of Y. This is reflected in the term $(\widehat{fh}_t - \hat{f}_t\hat{h}_t)\,dZ_t$. Note that $\widehat{fh}_t - \hat{f}_t\hat{h}_t$ is the covariance of f_t and h_t, conditional on $\mathcal{F}_t^Y$. The formula (4) generalizes the linear prediction formula
$$\hat{x} = \bar{x} + \frac{\mathrm{cov}(x,y)}{\mathrm{var}(y)}(y - \bar{y}),$$

which yields $\hat{x} = E[x|y]$ when x and y are joint normal.

We consider two applications.

2.1. **Kalman-Bucy Filter.** Assume X_0 is distributed normally with variance σ^2 and

$$\begin{aligned} dX_t &= aX_t\,dt + dB_t, \\ dY_t &= cX_t\,dt + dW_t, \end{aligned}$$

where B and W are independent real-valued Brownian motions that are independent of X_0. In this case, the distribution of X_t conditional on $\mathcal{F}_t^Y$ is normal with deterministic variance v_t. Moreover,

$$d\hat{X}_t = a\hat{X}_t\,dt + cv_t\,dZ_t, \tag{2.4}$$

where the innovation process Z is given by

$$dZ_t = dY_t - c\hat{X}_t\,dt. \tag{2.5}$$

Furthermore,

$$v_t = \frac{\gamma\alpha e^{\lambda t} - \beta}{\gamma e^{\lambda t} + 1}, \tag{2.6}$$

where α and $-\beta$ are the two roots of the quadratic equation $1 + 2ax - c^2x^2 = 0$, with both α and β positive, $\lambda = c^2(\alpha + \beta)$ and $\gamma = (\sigma^2 + \beta)/(\alpha - \sigma^2)$. One can consult Rogers and Williams (2000) for the derivation of these results from the general filtering results cited above.

2.2. **Two-State Markov Chain.** A simple tractable model that lies outside the Gaussian family is a two-state Markov chain. It is convenient to label the states zero and one, so assume X takes values in $\{0, 1\}$. Assume

$$dX_t = (1 - X_{t-})\,dN_t^0 - X_{t-}\,dN_t^1,$$

where $X_{t-} \equiv \lim_{s\uparrow t} X_s$ and the N^i are independent Poisson processes with parameters λ^i that are independent of X_0. This means that X stays in each state an exponentially distributed amount of time, with the exponential distribution determining the transition from state i to state j having parameter λ^i. Assume

$$dY_t = \theta(X_{t-})\,dt + dW_t,$$

where W is a Brownian motion on $\mathbb{R}^n$ independent of the N^i and X_0. Thus, the drift vector of Y is $\theta(0)$ or $\theta(1)$ depending on the state X_{t-}. In terms of our earlier notation, $h_t = \theta(X_{t-})$, and obviously we have

$$h_t = (1 - X_{t-})\theta(0) + X_{t-}\theta(1).$$

Write π_t for $\hat{X}_t$. This is the conditional probability that $X_t = 1$. The general filtering formula implies

$$d\pi_t = \big[(1-\pi_t)\lambda^0 - \pi_t\lambda^1\big]\,dt + \pi_t(1-\pi_t)\big[\theta(1)-\theta(0)\big]'\,dZ_t, \tag{2.7}$$

where the innovation process Z is given by

$$dZ_t = dY_t - \big[(1-\pi_t)\theta(0) + \pi_t\theta(1)\big]\,dt. \tag{2.8}$$

This is a special case of the results in Liptser and Shiryayev (1977, Chapter 9).

Note the similarity of (2.7) with the Kalman-Bucy filter (2.4): $\theta(1)-\theta(0)$ is the vector c in the equation

$$\begin{aligned} dY_t &= \theta(X_{t-})\,dt + dW_t \\ &= \big[(1-X_{t-})\theta(0) + X_{t-}\theta(1)\big]\,dt + dW_t \\ &= \theta(0)\,dt + cX_{t-}\,dt + dW_t, \end{aligned}$$

and $\pi_t(1-\pi_t)$ is the variance of X_t conditional on $\mathcal{F}_t^Y$.

3. Incomplete Information

3.1. Seminal Work.

Seminal Work. Early work in portfolio choice and market equilibrium under incomplete information includes Detemple (1986), Dothan and Feldman (1986), and Gennotte (1986). They analyze models of the following sort. The instantaneous rate of return on an asset is given by

$$\frac{dS}{S} = \mu_t\,dt + \sigma\,dW,$$

where

$$d\mu_t = \kappa(\theta - \mu_t)\,dt + \phi\,dB$$

and W and B are Brownian motions with a constant correlation coefficient ρ, and where μ_0 is normally distributed and independent of W and B. It is assumed that investors observe S but not μ; i.e., their filtration is the (augmented) filtration generated by S. As before, letting $\hat{\mu}$ denote the optional projection of μ, the innovation process is

$$dZ = \frac{\mu_t - \hat{\mu}_t}{\sigma}\,dt + dW,$$

which is an $\{\mathcal{F}_t^S\}$–Brownian motion. Moreover, we can write

$$\frac{dS}{S} = \hat{\mu}_t\,dt + \sigma\,dZ. \tag{3.1}$$

Because $\hat{\mu}$ is observable (adapted to $\{\mathcal{F}_t^S\}$), this is equivalent to a standard complete information model, and the portfolio choice theory of Merton applies to (3.1). This is a particular application of the separation principle for optimal control under incomplete information, and in fact the primary contribution of these early papers was to highlight the role of the separation principle. It is worthwhile to point out that in these models, it is $\hat{\mu}$ rather than μ that should be interpreted as the expected return on the asset, because the term "expected return" should represent an agent's beliefs, conditional on his information, and the beliefs of an agent are by definition observable to him.

These early models were interpreted as equilibrium models by assuming the returns are the returns of physical investment technologies having constant returns to scale, as in Cox, Ingersoll and Ross (1985). In other words, the assets are in infinitely elastic supply. In this case, there are no market clearing conditions to be satisfied. Equilibrium is determined by the optimal investments and consumption of the agents. Given an equilibrium, prices of other zero net supply assets can be determined—for example, term structure models can be developed. However, the set of such models that can be generated by assuming incomplete information is the same as the set that can be generated with complete information, given the equivalence of (3.1) and complete information models. In particular, the Kalman-Bucy filtering equations imply particular dynamics for $\hat{\mu}$, but one could equally well assume the same dynamics for μ and assume μ is observable.

3.2. **Markov Chain Models.** In Gaussian models (with Gaussian priors) the conditional covariance matrix of the unobserved variables is deterministic. This means that there is no real linkage between Gaussian incomplete information models and the well-documented phenomenon of stochastic volatility. Detemple (1991) observes that within a model that is otherwise Gaussian, stochastic volatility can be generated by assuming non-Gaussian priors. However, more recent work, motivated by issues regarding stochastic volatility, has focused on Markov chain models.

David (1993, 1997) studies a Cox-Ingersoll-Ross economy, assuming a two-state Markov chain, with an exponential distribution for the transition time from each state, as in Section 2.2.[1] In his model, there are two assets ($i = 0, 1$), with

$$\frac{dS^i}{S^i} = \mu^i(X_{t-})\,dt + \sigma^i\,dW^i,$$

[1]See also Honda (1997), though much of Honda's work seems to have been anticipated by David (1993).

where W^0 and W^1 are independent Brownian motions, $X_t \in \{0,1\}$, and $\mu^0(x) = \mu^1(1-x)$. Set $\mu_a = \mu^0(0)$ and $\mu_b = \mu^0(1)$. Then when $X_{t-} = 0$, the growth rates of the assets are μ_a for asset 0 and μ_b for asset 1, and the growth rates of the assets are reversed when $X_{t-} = 1$. With complete information in this economy, the investment opportunity set is independent of X_{t-}. However, with incomplete information, investors do not know for certain which asset is most productive. Suppose, for example, that $\mu_a > \mu_b$. Then asset 0 is most productive in state 0 and asset 1 is most productive in state 1. The filtering equation for the model is (2.7), with observation process $Y = (Y^0, Y^1)$, where

$$dY_t^i = \frac{d\log S_t^i}{\sigma^i} = \left(\frac{\mu^i(X_{t-})}{\sigma^i} - \frac{\sigma^i}{2}\right) dt + dW^i.$$

In terms of the innovation processes (the following equations actually define the innovation processes), we have

$$\begin{aligned} \frac{dS^0}{S^0} &= \left[(1-\pi_t)\mu_a + \pi_t\mu_b\right] dt + \sigma^0\, dZ^0, \\ \frac{dS^1}{S^1} &= \left[\pi_t\mu_a + (1-\pi_t)\mu_b\right] dt + \sigma^1\, dZ^1. \end{aligned}$$

As in the work of Detemple (1986), Dothan–Feldman (1986), and Gennotte (1986), this is equivalent to a complete information model in which the expected rates of return of the assets are stochastic with particular dynamics given by the filtering equations, but the volatilities of assets are constant.

David focuses on the volatility of the market portfolio, assuming a representative investor with power utility. The weights of the two assets in the market portfolio will depend on π_t (e.g., asset 0 will be weighted more highly when π_t is small, because this means a greater belief that the expected return of asset 0 is $\mu_a > \mu_b$). Due to diversification, the instantaneous volatility of the market portfolio will be smallest when the assets are equally weighted, which will be the case when $\pi_t = 1/2$, and the volatility will be higher when π_t is near 0 or 1. Therefore, the market portfolio will have a stochastic volatility. Using simulation evidence, David shows that the return on the market portfolio in the model is consistent with the following stylized facts regarding asset returns.

1) Excess kurtosis: the tails of asset return distributions are "too fat" to be consistent with normality.
2) Skewness: large negative returns occur more frequently than large positive returns.

3) Covariation between returns and changes in conditional variances: large negative returns are associated with a greater increase in the conditional variance than are large positive returns.

Arguably, a more interesting context in which to study incomplete information is an economy of the type studied by Lucas (1978), in which the assets are in fixed supply. In this case, the prices and returns of the assets are determined in equilibrium by the market-clearing conditions and hence will be affected fundamentally by the nature of information.

Veronesi (1999, 2000) and David and Veronesi (2002) study models of this type and discuss various issues regarding the volatility and expected return of the market portfolio. Their models are variations on the following basic model. Assume there is a single asset, with supply normalized to one, which pays dividends at rate D. Assume

$$\frac{dD_t}{D_t} = \alpha_D(X_{t-})\,dt + \sigma_D\,dW^1, \tag{3.2}$$

where X is a two-state Markov chain with switching between states occurring at exponentially distributed times, as in Section 2.2. Here W^1 is a real-valued Brownian motion independent of X_0. Investors observe the dividend rate D but do not observe the state X_{t-}, which determines the growth rate of dividends. We may also assume investors observe a process

$$dH_t = \alpha_H(X_{t-})\,dt + \sigma_H\,dW^2, \tag{3.3}$$

where W^2 is a real-valued Brownian motion independent of W^1 and X_0. The process H summarizes any other information investors may have about the state of the economy.

The filtering equations for this model are the same as those described earlier, where we set

$$Y = \left(\frac{\log D}{\sigma_D}, \frac{H}{\sigma_H}\right) \quad \text{and} \quad \mu = \left(\frac{\alpha_D - \sigma_D^2/2}{\sigma_D}, \frac{\alpha_H}{\sigma_H}\right).$$

In terms of the innovation process $Z = (Z^1, Z^2)$, we have

$$\frac{dD_t}{D_t} = \left[\pi_t\alpha_D(1) + (1-\pi_t)\alpha_D(0)\right]dt + \sigma_D\,dZ^1, \tag{3.4}$$

$$dH = \left[\pi_t\alpha_H(1) + (1-\pi_t)\alpha_H(0)\right]dt + \sigma_H\,dZ^2, \tag{3.5}$$

and the conditional probability π_t evolves as

$$\begin{aligned} d\pi_t &= \left[(1-\pi_t)\lambda^0 - \pi_t\lambda^1\right]dt \\ &\quad + \pi_t(1-\pi_t)\left[\frac{\alpha_D(1)-\alpha_D(0)}{\sigma_D}\,dZ^1 + \frac{\alpha_H(1)-\alpha_H(0)}{\sigma_H}\,dZ^2\right]. \end{aligned} \tag{3.6}$$

Note that (3.4) and (3.6) form a Markovian system in which the growth rate of dividends is stochastic. From here, the analysis is entirely standard. It is assumed that there is a representative investor who is infinitely-lived and who maximizes the expected discounted utility of consumption $u(c_t)$, with discount rate δ. The representative investor must consume the aggregate dividend in equilibrium, and the price of the asset is determined by his marginal rate of substitution. Specifically, the asset price at time t must be

$$S_t = E\left[\int_t^\infty \frac{e^{-\delta(s-t)}u'(D_s)}{u'(D_t)} D_s\, ds \,\middle|\, \pi_t, D_t\right].$$

In the case of logarithmic utility, we obtain $S_t = D_t/\delta$, so the asset return is given by

$$\frac{dS_t}{S_t} = \big[\pi_t\alpha_D(1) + (1-\pi_t)\alpha_D(0)\big]\, dt + \sigma_D\, dZ^1.$$

This is essentially the same as the early models on incomplete information, because we have simply specified the expected return

$$\pi_t\alpha_D(1) + (1-\pi_t)\alpha_D(0)$$

as a particular stochastic process.

The case of power utility $u(c) = c^\gamma/\gamma$ is more interesting. In this case, we have

$$\begin{aligned}
S_t &= D_t^{1-\gamma} E\left[\int_t^\infty e^{-\delta(s-t)} D_s^\gamma\, ds \,\middle|\, \pi_t, D_t\right] \\
&= D_t^{1-\gamma}\left\{(1-\pi_t)E\left[\int_t^\infty e^{-\delta(s-t)} D_s^\gamma\, ds \,\middle|\, X_{t-} = 0, D_t\right]\right. \\
&\qquad \left. + \pi_t E\left[\int_t^\infty e^{-\delta(s-t)} D_s^\gamma\, ds \,\middle|\, X_{t-} = 1, D_t\right]\right\} \\
&= D_t\left\{(1-\pi_t)E\left[\int_0^\infty e^{-\delta t} e^{\gamma\int_0^t \left\{\left[\alpha_D(X_{t-}) - \sigma_d^2/2\right] ds + \sigma_D\, dZ_s^1\right\}}\, dt \,\middle|\, X_0 = 0\right]\right. \\
&\qquad \left. + \pi_t E\left[\int_0^\infty e^{-\delta t} e^{\gamma\int_0^t \left\{\left[\alpha_D(X_{t-}) - \sigma_d^2/2\right] ds + \sigma_D\, dZ_s^1\right\}}\, dt \,\middle|\, X_0 = 1\right]\right\},
\end{aligned}$$

which we can write as

$$D_t\big\{(1-\pi_t)C^0 + \pi_t C^1\big\},$$

for constants C^0 and C^1.

We obtain

$$\begin{aligned}
\frac{dS}{S} &= \frac{dD}{D} + \frac{(C^1 - C^0)\,d\pi}{(1-\pi)C^0 + \pi C^1} + \frac{(C^1 - C^0)\,d\langle D, \pi\rangle}{D[(1-\pi)C^0 + \pi C^1]} \\
&= \text{something}\,dt + \sigma_D\,dZ^1 \\
&\quad + \left[\frac{(C^1 - C^0)\pi(1-\pi)}{(1-\pi)C^0 + \pi C^1}\right] \\
&\qquad \times \left[\frac{\alpha_D(1) - \alpha_D(0)}{\sigma_D}\,dZ^1 + \frac{\alpha_S(1) - \alpha_S(0)}{\sigma_S}\,dZ^2\right].
\end{aligned} \tag{3.7}$$

The factor

$$\frac{(C^1 - C^0)\pi(1-\pi)}{(1-\pi)C^0 + \pi C^1} \tag{3.8}$$

introduces stochastic volatility. Thus, stochastic volatility can arise in a model in which the volatility of dividends is constant.

There are obviously other ways than incomplete information to introduce a stochastic growth rate of dividends in a Markovian model similar to (3.4) and (3.6). However, this approach leads to a very sensible connection between investors' uncertainty about the state of the economy and the volatility of assets. Note that the factor $\pi_t(1-\pi_t)$ in the numerator of (3.8) is the conditional variance of X_t—it is largest when π_t is near 1/2, when investors are most uncertain about the state of the economy, and smallest when π_t is near zero or one, which is when investors are most confident about the state of the economy. Thus, the volatility of the asset is linked to investors' confidence about future economic growth. However, it should kept in mind that the model only characterizes the market portfolio. Yan (2003) argues that the qualitative aspects of the model do not apply to individual assets.

Veronesi (1999) actually assumes that the level of dividends (rather than the logarithm of dividends) follows an Ornstein-Uhlenbeck process as in (3.2) and he assumes the representative investor has negative exponential utility (i.e., he assumes constant *absolute* risk aversion rather than constant *relative* risk aversion). David and Veronesi (2002) study the model described here but assume the representative investor also has an endowment stream. They show that the model can generate a time-varying correlation between the return and volatility of the market portfolio (for example, sometimes the correlation may be positive and sometimes it may be negative) and use the model to generate an option pricing formula for options on the market portfolio.

Veronesi (2000) studies the above model but assuming there are n states of the world rather than just two. One way to express his model is to let the state variable X_t take values in $\{1, \ldots, n\}$ with dynamics

$$dX_t = \sum_{i=1}^{n} (i - X_{t-})\, dN_t^i,$$

where the N^i are independent Poisson processes with parameters λ^i. Then $N \equiv \sum_{i=1}^n N^i$ is a Poisson process with parameter $\lambda \equiv \sum_{i=1}^n \lambda^i$. Conditional on $\Delta N_t = 1$, there is probability $f^i = \lambda^i/\lambda$ that $X_t = i$. This is independent of the prior state X_{t-}. Define $X_t^i = 1_{\{X_t=i\}}$. Then the $\{\mathcal{F}_t^Y\}$–optional projection of X_t^i, which we will denote by π_t^i, is the probability that $X_t = i$ conditional on $\mathcal{F}_t^Y$. The process X_t^i is a two-state Markov chain with dynamics

$$dX_t^i = (1 - X_{t-}^i)\, dN_t^{-i} - X_{t-}^i\, dN_t^i,$$

where $N^{-i} \equiv \sum_{j \neq i} N_t^j$ is a Poisson process with parameter $\lambda^{-i} \equiv \sum_{j \neq i} \lambda^j$. Thus, the dynamics of π^i are given by the filtering equation (2.7) for two-state Markov chains. The resulting formula for the dynamics of the asset price S is a straightforward generalization of (3.7).

3.3. **Heterogeneous Priors.** Economists often assume that all agents have the same prior beliefs. A rationale for this assumption is given by Harsanyi (1967). To some, this rationale seems less than compelling, motivating the analysis of heterogeneous prior beliefs. A good example is Detemple-Murthy (1997). They study a single-asset Lucas economy similar to the one described in the previous section (but with the unobservable dividend growth rate being driven by a Brownian motion instead of following a two-state Markov chain). Instead of assuming a representative investor, they assume there are two classes of investors with different beliefs about the initial value of the dividend growth rate. Finally, they assume each type of investor has logarithmic utility and the investors all have the same discount rate. The focus of their paper is the impact of margin requirements, which limit short sales of the asset and limit borrowing to buy the asset. This is an example of an issue that cannot be addressed in a representative investor model, because margin requirements are never binding in equilibrium on a representative investor, given that he simply holds the market portfolio in equilibrium. In a frictionless complete-markets economy one can always construct a representative investor, but that is not necessarily true in an economy with margin requirements or other frictions or incompleteness of markets. In the absence of a representative investor, it can be difficult to compute or characterize an equilibrium, but this task is considerably simplified by assuming logarithmic utility, because that implies investors are “myopic”—they hold the tangency portfolio and do not have hedging demands. However, if all investors have logarithmic utility, then heterogeneity must be introduced through some other mechanism than the utility function. The assumption of incomplete information and heterogeneous priors is a simple device for generating this heterogeneity among agents. Basak and Croitoru (2003) study the effect of introducing "arbitrageurs" (for example, financial intermediaries) in the model of Detemple and Murthy. Jorion and Napp (2003) discuss the existence of representative investors in markets with incomplete information and heterogeneous beliefs.

4. Asymmetric Information

4.1. **Non-Equilibrium Models.** Recently, a literature has developed using the theory of enlargement of filtrations to study the topic of "insider trading." See, for example, Karatzas-Pikovsky (1996), Grorud-Pontier (1998, 2001), and Baudoin (2003). One starts with asset prices of the usual form[2]

$$\frac{dS_t^i}{S_t^i} = \mu_t^i\,dt + \sigma_t^i\,dW_t^i,$$

on the horizon $[0,T]$ where the W^i are correlated Brownian motions on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$. Then one supposes there is an $\mathcal{F}_T$–measurable random variable Y (with values in $\mathbb{R}^k$ or some more general space) and an "insider" has access to the filtration $\{\mathcal{G}_t\}$, which is the right-continuous completed version of the filtration $\{\mathcal{F}_t \vee \sigma{-}(Y)\}$. By "access to the filtration," I mean that the insider is allowed to choose trading strategies that are $\{\mathcal{G}_t\}$–adapted.

Some interesting questions are (1) does the model make mathematical sense—i.e., are the price processes $\{\mathcal{G}_t\}$–semimartingales? (2) is there an arbitrage opportunity for the insider? (3) is the market complete for the insider? (4) how much additional utility can the insider earn from his advance knowledge of Y? (5) how would the insider value derivatives? .... For the answer to the first question, the essential reference is Jacod (1985). Baudoin (2003) describes the setup I have outlined here as the case of "strong information" and also introduces a concept of "weak information."

While this literature is quite interesting in some respects, an essential limitation is that there is no feedback from the inside information of the investor to the dynamics of market prices. Suppose for example that there is a constant riskless rate r and the advance information Y

[2] Assume either that there are no dividends or that the S_i represent the prices of the portfolios in which dividends are reinvested in new shares.

is the vector of asset prices S_T. Then there is an arbitrage opportunity for the insider unless

$$S_t^i = e^{-r(T-t)} S_T^i$$

for all i and t, which of course cannot be the case if the volatilities σ^i are nonzero. One might simply say that this is not an acceptable model and adopt hypotheses that exclude it. However, the rationale for excluding it must be that we do not believe in the presence of arbitrage opportunities, and the rationale for that belief must be that exploitation of arbitrage opportunities tends to eliminate them. In other words, buying and selling by the insider would be expected to change market prices. This is true in general and not just in this specific example. The idea that market prices reflect in some way and to some extent the information of economic agents is a cornerstone of finance and of economics in general, dating back at least to Hayek (1940s?). In the following sections, we will examine equilibrium models, which incorporate this feedback from information to prices.

4.2. **Rational Expectations Models.** The term "rational expectations" was originally used in the context of asymmetric information models to mean that agents understand the mapping from the information of various agents to the equilibrium price; thus they make correct inferences from prices (see Grossman (1981)). These models were "competitive" models in the sense that agents were assumed to be "price takers," meaning that they assume their own actions have no effect on prices. Now the term is generally reserved for competitive models, and I will use it in that sense. We will examine strategic models, in which agents understand the impact of their actions on prices, in the next section.

An important rational expectations model is that of Wang (1993). Wang studies a Lucas economy in which the dividend rate D_t of the asset has dynamics

$$dD_t = (\Pi_t - kD_t)\,dt + b_D\,dW, \tag{4.1}$$

where W is an $\mathbb{R}^3$–valued Brownian motion. Moreover, it is assumed that

$$d\Pi_t = a_\Pi(\bar{\Pi} - \Pi_t)\,dt + b_\Pi\,dW \tag{4.2}$$

for a constant $\bar{\Pi}$. It is also assumed that there is a Cox-Ingersoll-Ross-type asset (i.e., one in infinitely elastic supply) that pays the constant rate of return r. There are two classes of investors, each having constant absolute risk aversion.

One class of investors (the "informed traders") observes D and Π. The other class (the "uninformed traders") observes only D. As described thus far, the model should admit a "fully revealing equilibrium," in which the uninformed traders could infer the value of Π_t from the equilibrium price of the asset. This equilibrium suffers from the "Grossman-Stiglitz paradox"—in reality it presumably costs some effort or money to become informed, but if prices are fully revealing, then no one would pay the cost of becoming informed; however, if no one is informed, prices cannot be fully revealing (and it would presumably be worthwhile in that case for someone to pay the cost of becoming informed). Wang avoids this outcome by the device introduced by Grossman and Stiglitz (1980): he assumes the asset is subject to supply shocks that are unobserved by all traders. The noise introduced by the supply shocks prevents uninformed traders from inverting the price to compute the information Π_t of informed traders.[3] Specifically, Wang assumes the supply of the asset is $1 + \Theta_t$, where

$$d\Theta_t = -a_\Theta \Theta \, dt + b_\Theta \, dW. \tag{4.3}$$

The general method used to solve rational expectations models is still that described by Grossman (1981), even though Grossman did not assume there were supply shocks and obtained a fully revealing equilibrium. The trick is to consider an "artificial economy" in which traders are endowed with certain additional information. One computes an equilibrium of the artificial economy and then shows that prices in this artificial economy reveal exactly the additional information traders were assumed to possess. Thus, the equilibrium of the artificial economy is an equilibrium of the actual economy in which traders make correct inferences from prices.

In Wang's artificial economy, the informed traders observe Θ as well as D and Π. The uninformed traders observe a linear combination of Θ and Π as well as D. In the equilibrium of the artificial economy, the price reveals the linear combination of Θ and Π, given knowledge of D. This implies that it reveals Θ to the informed traders, given that they are endowed with knowledge of Π and D. Therefore, the equilibrium of the artificial economy is an equilibrium of the actual economy.

Specifically, Wang conjectures that the equilibrium price S_t is a linear combination of D_t, Π_t, Θ_t and $\hat{\Pi}$, where $\hat{\Pi}$ denotes the optional projection of Π on the filtration of the uninformed traders. For this to

[3] In fact, this type of mechanism was first introduced by Lucas (1972), who assumes the money supply is unobservable in the short run, and hence real economic shocks cannot be intertwined from monetary shocks, leading to real effects of monetary policy in the short run.

make sense, one has to specify the filtration of the uninformed traders, and in the artificial economy it is specified as the filtration generated by D and a particular linear combination of Π and Θ. Let this linear combination be

$$H_t = \alpha\Pi_t + \beta\Theta_t \tag{4.4}$$

Then the observation process of the uninformed traders in the artificial economy is $Y_t = (D, H)$ and the unobserved process they wish to estimate is Π. For the equilibrium of the artificial economy to be an equilibrium of the actual economy, we will need S_t to be a linear combination of D_t, H_t and $\hat{\Pi}_t$; i.e.,

$$S_t = \delta + \gamma D_t + \kappa H_t + \lambda\hat{\Pi}_t. \tag{4.5}$$

Conditional on $\mathcal{F}_t^Y$, Π_t is normally distributed with mean $\hat{\Pi}_t$ and a deterministic variance. Wang derives an equilbrium in which S_t is a linear combination of D_t, Π_t, Θ_t and $\hat{\Pi}_t$ with time-invariant coefficients by focusing on the steady-state solution of the model. Specifically, he assumes the variance of Π_0 is the limit of the conditional variance of Π_t as $t \to \infty$.

Given the specification of the price process (4.4)–(4.5) and the filtering formula, it is straightforward to calculate the demands of the two classes of traders. The market clearing equation is that the sum of the demands equals Θ_t. This is a linear equation that must hold for all values of D_t, Π_t, Θ_t and $\hat{\Pi}_t$. Imposing this condition gives the equilibrium values of α, β, δ, γ, κ and λ.

In addition to the usual issues regarding the expected return and volatility of the market portfolio, Wang is able to describe the portfolio behavior of the two classes of investors; in particular, uninformed traders tend to act as "trend chasers," buying the asset when its price increases, and informed traders act as "contrarians," selling the asset when its price increases.

4.3. **Strategic Models.** The price-taking assumption in rational expectatons models is often problematic. In the extreme case, prices are fully revealing, and traders can form their demands as functions of the fully revealing prices, ignoring the information they possessed prior to observing prices. But, if traders all act independently of their own information, how can prices reveal information? (reference Beja (1980?)). Moreover, as mentioned earlier, full revelation of information by prices would eliminate the incentive to collect information in the first place.

The price-taking assumption is particularly problematic when information is possessed by only one or a few traders. Consider the case of a piece of information that is held by only a single trader. In general, the equilbirium price in a rational expectatons model will reflect this information to some extent. Moreover, traders are assumed to make correct inferences from prices, so the trader is assumed to be aware that his information enters prices. But how can he anticipate that the price will reflect his private information, when he assumes that his actions do not affect the price? Hellwig (1980) describes this as "schizophrenia" on the part of traders.

These issues do not arise in strategic models, in which agents are assumed to recognize that their actions affect prices and it is only through their actions that private information becomes incorporated into prices. The most prominent model of strategic trading with asymmetric information is due to Kyle (1985). Kyle's model has been applied on many occasions, beginning with Admati–Pfleiderer (1988), to study various issues in market microstructure. The continuous-time version of the model was formalized and generalized by Back (1992). Subsequent generalizations appear in Back (1993), Back-Pedersen (1998), Baruch (2002), Cho (2003), and Lasserre (forthcoming). Here we will discuss the model with multiple informed traders due to Back, Cao and Willard (2000). Their work builds on the analysis by Foster and Viswanathan (1996) of a discrete-time model with multiple traders. We will also discuss the connection derived by Back and Baruch (2003) between the Kyle model and the Glosten-Milgrom (1985) model .

The Kyle model focuses on a single risky asset traded over the time period $[0, T]$. It is assumed that there is also a riskless asset, with the risk-free rate normalized to zero. Unlike models described previously in which the single risky asset is interpreted as the market portfolio, with the dividend of the asset equaling aggregate consumption, the Kyle model is not a model of the market portfolio. In fact, the risk of the asset is best interpreted as idiosyncratic, because investors are assumed to be risk neutral. As in Grossman-Stiglitz (1980) or Wang (1993) it is assumed that the supply of the asset is subject to random shocks, which we interpret as resulting from the trade of "noise traders." The noise traders trade for reasons that are unmodeled. For example, they may experience liquidity shocks (endowments of cash to be invested or desires for cash for consumption) and for that reason are often called "liquidity traders." In addition to N strategic traders and the noise traders, it is assumed that there are competitive risk-neutral "market makers," who are somewhat analogous to the uninformed traders in Wang (1993). The market makers observe the net demands of the strategic traders and noise traders and compete to fill their demands. As a result of their competition (and their risk neutrality and the fact that the risk-free rate is zero), the transaction price is always the expectation of the asset value, conditional on the information of the market makers, i.e., conditional on the information in the history of orders.

In the original model of Kyle (1985), there is only a single strategic trader. He possesses at date zero some information about the value of the asset at date T that other traders do not have. Given some specification for how his trades (and hence how the net trades of the informed and noise traders) depend on his information, the market makers filter the order flow to estimate it. In the multiple strategic trader model of Back, Cao and Willard (2000), each strategic trader must also filter the order flow to estimate the information of other strategic traders.

It is assumed that the information asymmetry is erased by a public announcement at date T. Since this eliminates the "lemons problem," all positions can be liquidated at this announced value. Denote this value by v.

Before beginning the discussion of the model with multiple strategic traders, it will be useful to first explicate the single-trader model of Kyle. In fact, as in Kyle (1985), we will begin with a single-period model. In this model, there is trading only at date 0, and consumption occurs at date T. The asset value v is normally distributed with mean $\bar{v}$ and variance σ_v^2. The informed trader observes v and submits an order $x(v)$. Noise traders submit an order z that is independent of v and normally distributed with mean zero and variance σ_z^2. Market makers observe $y \equiv x + z$ and set the price equal to $p = E[v|x(v) + z]$. We search for a linear equilibrium in which the price is set as $p = \bar{v} + \lambda y$ and the insider's trade is $x = \eta(v - \bar{v})$, for constants λ and η. An equilibrium is defined by (1) $\bar{v} + \lambda y = E[v|y]$ and (2) $\eta(v - \bar{v}) = \text{argmax}_x \, E[x(v - \bar{v} - \lambda(x + z))]$. Condition (1) implies

$$\lambda = \frac{\text{cov}(v, y)}{\text{var}(y)} = \frac{\eta \sigma_v^2}{\eta^2 \sigma_v^2 + \sigma_z^2},$$

and condition (2) implies

$$\eta = \frac{1}{2\lambda}.$$

The solution of these two equations is

$$\eta = \frac{\sigma_z}{\sigma_v} \qquad \text{and} \qquad \lambda = \frac{\sigma_v}{2\sigma_z}.$$

Kyle defines the reciprocal of λ as the "depth" of the market. It measures the number of shares that can be traded causing only a unit change in the price. Of interest is the fact that the depth of the market is proportional to the amount of noise trading as measured by σ_v and inversely proportional to the amount of private information as measured by σ_v.

To be continued ...

References

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

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

Back, K. (1993): "Asymmetric Information and Options," *Review of Financial Studies*, 6, 435–472.

Back, K., and S. Baruch (2003): "Information in Securities Markets: Kyle Meets Glosten and Milgrom," working paper, Washington University in St. Louis.

Back, K., and H. Pedersen (1998): "Long-Lived Information and Intraday Patterns," 1, 385–402.

Back, K., Cao, H., and G. Willard (2000): "Imperfect Competition among Informed Traders," *Journal of Finance*, 55, 2117–2155.

Baruch, S. (2002): "Insider Trading and Risk Aversion," *Journal of Financial Markets*, 5, 451–464.

Basak, S., and B. Croitoru (2003): "On the Role of Arbitrageurs in Rational Markets," preprint.

Baudoin, F. (2003): "The Financial Value of Weak Information on a Financial Market," in *Paris-Princeton Lectures on Mathematical Finance 2002*, Springer. next Brunnermeier, M. K. (2001): *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding*, Oxford University Press.

Cho, K.-H. (2003): "Continuous Auctions and Insider Trading: Uniqueness and Risk Aversion," *Finance and Stochastics*, 7, 47–71.

Cox, J., Ingersoll, J., and S. Ross, (1985): "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363–384.

David, A. (1993): "Business Cycle Risk and the Equity Premium," Ph.D. dissertation, University of California at Los Angeles.

David, A. (1997): "Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility," *Journal of Financial and Quantitative Analysis*, 32, 427–462.

David, A., and P. Veronesi (2002): "Option Prices with Uncertain Fundamentals: Theory and Evidence on the Dynamics of Implied Volatilities," preprint.

Detemple, J. (1986): "Asset Pricing in a Production Economy with Incomplete Information," *Journal of Finance*, 41, 383–391.

Detemple, J. (1991): "Further Results on Asset Pricing with Incomplete Information," *Journal of Economic dynamics and Control*, 15, 425–454.

Detemple, J., and S. Murthy (1997): "Equilibrium Asset Prices and No-Arbitrage with Portfolio Constraints," *Review of Financial Studies*, 10, 1133–1174.

Dothan, M. U., and D. Feldman (1986): "Equilibrium Interst Rates and Multiperiod Bonds in a Partially Observable Economy," *Journal of Finance*, 41, 369–382.

Foster, F. D., and S. Viswanathan (1996): "Strategic Trading when Agents Forecast the Forecasts of Others," *Journal of Finance*, 51, 1437–1478.

Fujisaki, M., Kallianpur, G., and H. Kunita (1972): "Stochastic Differential Equations for the Non-Linear Filtering Problem," *Osaka Journal of Mathematics*, 9, 19–40.

Gennotte, G. (1986): "Optimal Portfolio Choice under Incomplete Information," *Journal of Finance*, 41, 733–746.

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.

Grorud, A., and M. Pontier (1998): "Insider Trading in a Continuous Time Market Model," *International Journal of Theoretical and Applied Finance*, 1, 331–347.

Grorud, A., and M. Pontier (2001): "Asymmetrical Information and Incomplete Markets," *International Journal of Theoretical and Applied Finance*, 4, 285–302.

Grossman, S. (1981): "An Introduction to the Theory of Rational Expectations under Asymmetric Information," *Review of Economic Studies*, 31, 573–585.

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

Hellwig, M. (1980): "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 26, 279–312.

Honda, T. (1997): "Equilibrium Asset Pricing with Unobservable Regime-Switching Mean Earnings Growth," Stanford University.

Karatzas, I., and I. Pikovsky (1996): "Anticipative Portfolio Optimization," *Advances in Applied Probability*, 28, 1095–1122.

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

Jouni, E., and C. Napp (2003): "Consensus Consumer and Intertemporal Asset Pricing under Heterogeneous Beliefs," preprint.

Lasserre, G. (forthcoming): "Asymmetric Information and Imperfect Competition in a Continuous-Time Multivariate Security Model," *Finance and Stochastics*.

Lucas, R. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4, 103–124.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 40, 1429–1444.

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

VERONESI, P. (1999): "Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Model," *Review of Financial Studies*, 12, 975–1007.

VERONESI, P. (2000): "How Does Information Quality Affect Stock Returns," *Journal of Finance*, 55, 807–837.

WANG, J. (1993): "A Model of Intertemporal Asset Prices under Asymmetric Information," *Review of Economic Studies*, 60, 249–282.

YAN, H. (2003): "Information Quality and Asset Prices," preprint.

ZIEGLER, A. (2003): *Incomplete Information and Heterogeneous Beliefs in Continuous-Time Finance*, Springer.