The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

2229-1

School and Workshop on Market Microstructure: Design, Efficiency and Statistical Regularities

21 - 25 March 2011

Aggregation of Information and Rational Expectations

Marco LICALZI
Department of Management
University of Ca Foscari Venice
San Giobbe, Cannaregio 873, 30121
Venice
Italy

1. Aggregation of information and rational expectations

1.1 The normal-CARA world

1.1.1 Bayesian decision making under risk

Suppose that Primus is interested in taking actions that depend on an unknown parameter y. For instance, he has a utility function $u(y)$ which depends on the value of y. If Primus is a (subjective) expected utility maximizer, he should pick the action which maximizes the expected value of his utility computed with respect to his (subjective) probability distribution for y.

Suppose now that, before making any decision, Primus receives additional information about y. Then he should update his prior probability distribution to a *posterior* probability distribution and use this latter one to choose an optimal action. Updating beliefs when new information is available, therefore, is crucial for taking informed decisions.

This basic format for taking decisions under uncertainty when new information is revealed is the tenet of Bayesian rationality. Out of the many possible variations, the literature especially insists on a model where Primus has an exponential utility function and normally distributed random variables (both prior and posterior). This is made for analytical convenience. This section collects the basic mathematics necessary to deal with the *cara–normal* model.

1.1.2 Updating normal beliefs

Suppose that the prior distribution for the random variable Y is normal with mean m and standard deviation $s_y > 0$. For short, we write $Y \sim N(m, s_y)$. Imagine that Primus can receive signals about Y. Each signal x is independently and identically distributed according to a normal distribution with mean y and standard deviation $s_x > 0$; that is, a signal is an (iid) draw from $X \sim N(y, s_x)$.

We are interested in what should be Primus' posterior distribution for Y after having observed one signal X. If we denote by $g(y)$ the prior distribution for Y and by $f(x|y)$ the conditional distribution of the signal, we have respectively

$$g(y) = \frac{1}{\sqrt{2\pi} s_y} \exp\left[-\frac{1}{2s_y^2}(y-m)^2\right]$$

$$f(x|y) = \frac{1}{\sqrt{2\pi} s_x} \exp\left[-\frac{1}{2s_x^2}(x-y)^2\right]$$

By Bayes' rule, the posterior density function for Y given a signal $X = x$ is given by

$$g(y|x) = \frac{f(x|y) \cdot g(y)}{\int f(x|y) \cdot g(y)\, \mathrm{d}y}. \tag{1}$$

Carrying out substitutions, we can find the posterior density of $Y|x$ and check that

$$Y|x \sim N\left(\frac{\frac{m}{s_y^2} + \frac{x}{s_x^2}}{\frac{1}{s_y^2} + \frac{1}{s_x^2}}, \frac{1}{\frac{1}{s_y^2} + \frac{1}{s_x^2}} \right). \tag{2}$$

Three properties are worth being noted. First, the posterior is a normal as well. If we begin with a normal prior and the signal is normally distributed, the posterior remains normal. This feature is extensively used in models with rational expectations.

Second, we can simplify (2) by defining the *precision* of a normally distributed signal as the inverse of its variance. In particular, let $\tau_y = (1/s_y^2)$ and $\tau_x = (1/s_x^2)$ respectively the precisions of Y and X. Then (2) can be written as

$$Y|x \sim N\left(\frac{m\tau_y + x\tau_x}{\tau_y + \tau_x}, \frac{1}{\tau_y + \tau_x} \right). \tag{3}$$

Thus, the posterior mean of $Y|x$ can be written more simply as the average of the prior mean and of the signal weighted by their respective precisions. In the following, we make frequent use of this simple method for computing the expected value of a posterior belief.

Third, note that the Bayesian posterior beliefs converge to the truth as the number of signals increase. After n (iid) draws $x_1, x_2, \ldots, x_n$, the variance of the posterior goes to zero while the Strong Law of Large Numbers implies that the posterior mean converges to y.

1.1.3 Cara preferences in a normal world

If Primus is an expected utility maximizer with constant absolute risk aversion, his utility function must be linear or exponential. In particular, if we also assume that he is strictly risk averse, his utility function over the wealth w must be a negative exponential

$$u(w) = -e^{-kw} \tag{4}$$

where $k > 0$ is his coefficient of (absolute) risk aversion.

Suppose that Primus has preferences which satisfy these assumptions and that his beliefs are normally distributed so that $W \sim N(\mu, \sigma)$. Check that his expected utility can be written

$$Eu(W) = \int \left[-e^{-kw}\right] \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(w-\mu)^2\right] \mathrm{d}w = -\exp\left\{-\left[k\mu - \frac{1}{2}k^2\sigma^2\right]\right\}.$$

This expression is a function only of the current mean μ and the current variance σ^2 of W. Since $-\exp(-kw)$ is a strictly increasing function of w, Primus' preferences in a cara–normal world are more simply expressed by the functional

$$V(\mu, \sigma) = \mu - \frac{1}{2}k\sigma^2$$

and we can characterize them using simply the two statistics μ and σ^2 and the coefficient of absolute risk aversion k.

1.1.4 Demand for a risky asset

Suppose that Primus is an expected utility maximizer with constant absolute risk aversion $k > 0$. Assume that there are two assets, one of which is a risky stock and the other is a riskless bond. The bond has a current price normalized to 1 and will pay a riskless amount $(1 + r)$ at the end of the period. The stock has a current price of p and will pay a risky amount $Y \sim N(m, s)$ at the end of the period. Primus' current endowment is w. What is Primus' optimal portfolio?

Primus is interested in maximizing the expected value of his terminal wealth (at the end of the period). Assuming that short sales are allowed and the market is frictionless, he can invest in any portfolio of α stocks and β bonds such that $\alpha p + \beta = w$. Therefore, $\beta = w - \alpha p$. Note that α and β are unrestricted in sign and may not add to 1.

The terminal value of such a portfolio is normally distributed with mean $\alpha m + \beta(1+r) = \alpha m + (w - \alpha p)(1+r)$ and variance $\alpha^2 s^2$. Hence, the expected utility of Primus of a portfolio with α stocks is

$$\alpha m + (w - \alpha p)(1 + r) - \frac{1}{2} k \alpha^2 s^2.$$

Maximizing with respect to α, this yields Primus' demand function for stock:

$$\alpha(p, r; m, s^2; k) = \frac{m - p(1 + r)}{ks^2}. \tag{5}$$

Note that the demand for the risky stock does not depend on the initial wealth w. Moreover, the demand is separately monotone in each of its arguments; for instance, it is increasing in the mean m and decreasing in the variance s^2.

1.2 Rational expectations

1.2.1 Introduction

The main question addressed by rational expectations models is what happens when people with different information decide to trade. How market prices are affected by traders' information affects how the traders can infer information from market prices. The fundamental insight is that prices serve two purposes: they clear markets and they aggregate information. This dual role can make the behavior of prices and markets much more complex than assumed in simple models of asset behavior.

Let us begin with an example. Suppose that there are two agents in the market for q widgets. Primus receives a binary signal about the true value of widgets: if the signal is High, his demand for widgets is $p = 5 - q$; if the signal is Low, his demand is $p = 3 - q$. We say that Primus is *informed* because his demand depends on which signal he receives. Secunda receives no signal and offers an unconditional supply of widgets $p = 1+q$. Moreover, assume that, if she could receive signals, Secunda would change her supply to $p = 1 + 3q$ with an H-signal and to $p = 1$ with an L-signal.

When Secunda is sufficiently naive, the following situation occurs. If Primus receives an H-signal, the demand from the informed Primus equates the supply from an uninformed Secunda at a price of $p^H = 3$ (and $q = 2$ widgets are exchanged). If he receives an L-signal, his demand equates the supply from Secunda at a price of $p^L = 2$ (and $q = 1$ widget is exchanged). Different prices clear markets for different signals: $p = 3$ when the signal is H and $p = 2$ when it is L.

This outcomes, however, presumes that Secunda does not understand that prices also convey information. The market-clearing price is $p = 3$ if (and only if) the signal is H. Thus, if Secunda sees that markets clear at a price of $p = 3$ she can infer that Primus has received an H-signal and this suffices to let her change the supply function to $p = 1+3q$. But in this case the market must clear at a price such that $5-q = 1+3q$, that is $p = 4$. Similarly, if the market-clearing price would be $p^L = 2$, Secunda would understand that Primus got an L-signal and her supply would switch to $p = 1$, making this the market-clearing price.

In other words, if Secunda passively lets the prices clear the market, the prices are $p^H = 3$ and $p^L = 2$. If she exploits the information embedded in different prices, the prices will be $p^H = 4$ and $p^L = 1$. The first case ($p^H = 3$ and $p^L = 2$) can be an equilibrium only if we assume that Secunda is not sufficiently rational to understand that prices reveal information, or to use the information which is revealed by prices. Market-clearing equilibria with rational agents require that the information embedded in prices is fully exploited, and this is what the notion of rational expectations equilibrium is about.

1.2.2 A simple financial market

We consider a two-asset one-period economy in which all random variables are independent and normally distributed, with strictly positive standard deviations. The two available assets are a risky stock and a riskless bond. The bond has a current price normalized to 1 and will pay a riskless amount $(1+r)$ at the end of the period. The stock has a current price of p and will pay a risky amount $Y \sim N(m, s_y)$ at the end of the period. For convenience, denote by $\tau_y = 1/s_y^2$ the precision of Y.

There are n traders in the economy. They have identical cara preferences expressed by the utility function (4), defined over terminal wealth, with $k = 1$. Each agent i has an initial endowment w_i of wealth and receives a signal $X_i \sim N(y, s_x)$, with $s_x > 0$. Agents' signals are independent and identically distributed. The overall supply of stock is exogenously fixed to $z \geq 0$. (The specific value of z does not matter: most people like to normalize it to 1, but a few purists set it to 0 on the ground that the stock is created and exchanged among the market participants.)

Suppose that agents are naive and do not consider the informational value of the price. Based on the realization x_i of his own signal, agent i's demand for stock is

$$\alpha_i = m\tau_y + x_i\tau_x - p(1+r)(\tau_y + \tau_x). \tag{6}$$

It can be shown that (6) follows from (5) using the fact that the agent maximizes expected utility based on his *posterior* distribution.

The competitive equilibrium requires that the aggregate demand equates the aggregate supply; that is, $\sum_i a_i = z$, which gives the competitive equilibrium price

$$p = \frac{m\tau_y + \bar{x}\tau_x - (z/n)}{(1+r)(\tau_y + \tau_x)}, \tag{7}$$

where

$$\bar{x} = \frac{\sum_i x_i}{n}$$

is the average of all the agents' signals.

It is clear that (7) describes a 1-to-1 function between price and $\bar{x}$. Hence, given a price p, an agent can invert this function and learn the realization of the random variable $\bar{X}$. Since $\bar{X} \sim N(y, s_y/\sqrt{n})$, this can be viewed as a signal about Y that has a precision higher than that of X_i. Hence, the competitive price aggregates the information singularly received by each agent into a more powerful statistics.

Suppose now that an agent is not naive and recognizes that he can exploits the information aggregated by the price. Then he would change his demand to take into account the new information. Recall that the agent knows both his own signal X_i and (by inverting the price) $\bar{X}$. It can be shown that the distribution of Y conditional on $\bar{X}$ and X_i is the same as the distribution of Y conditional on $\bar{X}$. Hence, the updated posterior for agent i can be simply computed by assuming that the agent has a prior $Y \sim N(m, s_y)$ and receives a signal $\bar{X} \sim N(y, s_y/\sqrt{n})$. If we let $\tau_{\bar{x}}$ denote the precision of the signal $\bar{X}$ and $\bar{x}$ be the realization of the "signal" $\bar{X}$, the posterior distribution is

$$Y|\bar{X} = \bar{x} \sim N\left(\frac{m\tau_y + \bar{x}\tau_{\bar{x}}}{\tau_y + \tau_{\bar{x}}}, \frac{1}{\tau_y + \tau_{\bar{x}}} \right).$$

It follows that the demand of the agent would change to

$$\alpha_i = m\tau_y + \bar{x}\tau_{\bar{x}} - p(1+r)(\tau_y + \tau_{\bar{x}}), \tag{8}$$

which is different from the demand of the naive agent. This suggests that the competitive equilibrium with naive agents may not be stable to the recognition that, besides clearing the market, the price also aggregate information.

We are thus led to ask whether there exists an equilibrium when agents are not naive and instead have *rational expectations*. In a rational expectations equilibrium, the agents fully exploit the information revealed by the prices. It turns out that in our simple market the rational expectations equilibrium can be found very simply. Consider an artificial economy where every agent observes the same signal $\bar{X}$ and solve for the competitive equilibrium assuming naive agents. The demand of each agent is the same as in (8) and hence the equilibrium price of this artificial economy is

$$p = \frac{m\tau_y + \bar{x}\tau_{\bar{x}} - (z/n)}{(1+r)(\tau_y + \tau_{\bar{x}})}.$$

Inverting this price function still reveals the realization $\bar{x}$ of the signal $\bar{X}$. However, since now demands have been formed assuming knowledge of $\bar{x}$, an agent with rational expectations will not change his demands and therefore the price will not be affected.

This rational expectations equilibrium, however, may not be stable for a different reason. Note that agent i's demand function in (8) *does not* directly depend on his own information, but only on the statistics $\bar{x}$ which is identical for all agents. Now, if an agent's demand does not directly depend on his own information, how can the price aggregate individual information? Furthermore, if individual information does not enter an agent's demand, this agent has no incentive to collect information: he can let the others "do the job" of getting informed and then free-ride on them. But if all agents think this way or if there is a positive cost to collect information, how would the price formation mechanism get to aggregate information?

This problem arises (in part) because the rational expectations equilibrium just discussed reveals *completely* the information singularly received by each agents. We say that the equilibrium is *fully revealing*. Because the price reveals everything, there remain no incentives for an agent to collect information privately: everything he knows gets revealed to everybody for free. However, if we look at a partially revealing equilibrium with rational expectations, the price reveals only part of an agent's private information and therefore can still make it advantageous for him to search and collect private information (besides what the price might reveal).

1.2.3 Rational expectations equilibrium with noise

We consider a two-asset one-period economy in which all random variables are independent and normally distributed, with strictly positive standard deviations. The two available assets are a risky stock and a riskless bond. The bond has a current price normalized to 1 and will pay a riskless amount $(1+r)$ at the end of the period. The stock has a current price of p and will pay a risky amount $Y \sim N(m, s_y)$ at the end of the period. For convenience, denote by $\tau_y = 1/s_y^2$ the precision of Y.

There are two traders in the economy. (If you prefer, you can assume that there are two classes of traders and that within each class traders are identical.) They have identical cara preferences with $k = 1$ and arbitrary endowments of money. Primus is an informed trader, while Secunda is uninformed. Primus receives a signal $X \sim N(y, s_x)$ about the value of the stock; denote by $\tau_x = 1/s_x^2$ the precision of this signal. Secunda receives no signal. Differently from the previous example, we now assume that there is an exogenous and *random* supply of stock Z. For instance, this may come from "noise traders" who are forced to buy or sell stock at any price for external reasons such as a sudden and unexpected need for liquidity, or a bequest. For convenience, we assume that Primus and Secunda hold no stock and that $Z \sim N(0, s_z)$. Primus' signal is assumed to be independent of Z and thus neither trader receives information about Z.

By (3), Primus' posterior distribution is

$$Y|x \sim N\left(\frac{m\tau_y + x\tau_x}{\tau_y + \tau_x}, \frac{1}{\tau_y + \tau_x}\right).$$

Therefore, by (5), Primus' demand for the risky asset is

$$\alpha_1 = \frac{\frac{m\tau_y + x\tau_x}{\tau_y + \tau_x} - p(1+r)}{\frac{1}{\tau_y + \tau_x}} = m\tau_y + x\tau_x - p(1+r)(\tau_y + \tau_x).$$

Secunda, who is uninformed, receives no signal. However, she knows that trading from the informed agent affects the price. If she knew how this trading affects the market-clearing prices, she could extract information about the informed trader's signal from the market price.

A rational expectations equilibrium postulates that there is a specific *pricing rule* $P(\cdot)$ which links Primus' information with the market price and let Secunda use this rule to extract information. In equilibrium, the pricing rule must be correct; that is, Secunda must extract information that is consistent with Primus' signal.

Suppose that the pricing rule is linear (we will check this in a moment). That is, assume

$$p = am + bx - cz \tag{9}$$

for some appropriate coefficients a, b, c to be determined as part of the equilibrium. Given this pricing rule, Secunda (who is the uninformed trader) can construct the observable random variable

$$\eta := \frac{p - am}{b} = X - \frac{c}{b}Z.$$

Since $Z \sim N(0, s_z)$, this implies that η is an unbiased estimate of the signal X actually received by Primus. This estimate, however, is garbled by the additional zero-mean noise associated with Z. Since η is an unbiased estimator for X and X is an unbiased estimator for Y, η is also an unbiased estimator for Y. Indeed, as we see from the right-hand side, $\eta \sim N(y, s_x + (c/b)s_z)$. Therefore, η is a signal that Secunda can use to obtain information about Y.

Let τ_η be the precision associated with $s_\eta = s_x + (c/b)s_z$; of course, $\tau_\eta < \tau_x$. By (3), Secunda's posterior distribution for Y is

$$Y|\{P(\cdot), p\} \sim N\left(\frac{m\tau_y + \eta\tau_\eta}{\tau_y + \tau_\eta}, \frac{1}{\tau_y + \tau_\eta}\right).$$

Again by (5), Secunda's demand for the risky asset is

$$\alpha_2 = \frac{\frac{m\tau_y + \eta\tau_\eta}{\tau_y + \tau_\eta} - p(1+r)}{\frac{1}{\tau_y + \tau_\eta}} = m\tau_y + \eta\tau_\eta - p(1+r)(\tau_y + \tau_\eta).$$

The equilibrium price can be found by equating the aggregate demand $\alpha_1 + \alpha_2$ with the aggregate supply Z. This yields (after substituting for η)

$$p = \frac{2m\tau_y + x(\tau_x + \tau_\eta) - Z[1 + (c/b)\tau_\eta]}{(1+r)(2\tau_y + \tau_x + \tau_\eta)}.$$

This expression can be rewritten as $p = am + bx - cZ$ by appropriately choosing a, b, c such that

$$a = \frac{2\tau_y}{(1+r)(2\tau_y + \tau_x + \tau_\eta)},$$

$$b = \frac{\tau_x + \tau_\eta}{(1+r)(2\tau_y + \tau_x + \tau_\eta)},$$

$$c = \frac{1 + (c/b)\tau_\eta}{(1+r)(2\tau_y + \tau_x + \tau_\eta)}.$$

This confirms that the pricing rule conjectured in (9) is linear and concludes the example.

If the aggregate supply Z were not noisy, Secunda could use η to infer exactly what is the signal x received by Primus. We say that the rational expectations equilibrium is *fully revealing* if prices can be used to infer exactly what are the signals. In this case, prices are sufficient statistics for all the available information and the market is "efficient" in the strong sense (even private information is embedded in prices).

The rational expectations equilibrium is *partially revealing* if only a partial inference is possible, as in this example, where prices reflect both private information and exogenous noise. Now, if p goes up, Secunda cannot tell whether the cause is more positive private information (i.e., a higher signal for Primus) or smaller asset supply (i.e., a lower realization of Z).

1.3 Computing a rational expectations equilibrium

Suppose that there are n traders and j assets. Each trader observes a signal X_i ($i = 1, 2, \ldots, n$) about one or more assets. The construction of a rational expectations equilibrium (REE) can be outlined in five steps.

1. Specify each trader's prior beliefs and propose a pricing rule (which for the moment is only a conjecture) P^c mapping the traders' information to the prices of the assets. The pricing rule $P^c(X_1, X_2, \ldots, X_n, \varepsilon)$ may incorporate some noise ε. The traders takes this mapping as given. The pricing rule must be determined in equilibrium; at this stage, it is parameterized by undetermined coefficients because the true equilibrium price is not known yet.

2. Derive each trader's posterior beliefs, given the parameterized price conjectures and the important assumption that all traders draw inferences from prices. The posterior beliefs depend on the proposed pricing rule (e.g., from the undetermined coefficients).

3. Derive each trader's optimal demand, based on his (parameterized) beliefs and his preferences.

4. Impose the market clearing conditions for all markets and compute the endogenous market clearing prices. Since individual demands depend on traders' beliefs, so do prices. This gives the actual pricing rule $P^a(X_1, X_2, \ldots, X_n, \varepsilon)$ which provides the actual relationship between traders' signals and the prices.

5. Impose rational expectations; that is, make sure that the conjectured pricing rule P^c coincides with the actual pricing rule P^a. This can be achieved by equating the undetermined coefficients of P^c with the actual P^a.

1.4 An assessment of the rational expectations model

The REE provides a few key insights on which the following literature has built upon.

1. Prices play two roles: they clear markets and they convey information.

2. In a fully-revealing equilibrium, individual asset demands depend only on price, not on the trader's private information.

3. Therefore, in a fully-revealing equilibrium there is no incentive to invest in costly information: this incentive is restored in a partially-revealing equilibrium.

On the other sides, there are three important difficulties with the notion of a REE. First, there is the issue of its existence. If the number of possible signals is finite, then for a generic set of economies there exists a REE. This result, of course, does not apply to signals drawn from a normal distribution. Similarly, for j the number of assets in the economy, if the dimension S of the space of signals is lower than $j-1$, then for a generic set of economies there exists a REE. Intuitively, if there are more prices than signals, there is sufficient flexibility to both clear markets and aggregate information. On the other hand, if $S = j-1$, there is an open set of economies for which no REE exists. For instance, in a model with two assets (a risky stock and a riskless bond) and a one-dimensional signal, the existence of a REE, while possible, is a fragile result. Finally, for $S > j-1$, all weird things can happen.

Second, there is the issue of how the pricing rule is discovered. In a REE, traders must know the pricing rule that specifies the equilibrium prices. How such knowledge comes about, however, is left unspecified. One possible explanation is that it is learned over time, but learning to form rational expectations is not an easy task. During the learning process, traders must act according to "wrong" conjectures; their behavior, then, is likely to upset the emergence of the "correct" conjectures.

Third, there is the issue of price formation. In a REE, it is implicitly assumed that prices are set by a Walrasian auctioneer which (magically) equates demand and supply. This implicit auctioneer collects the "preliminary orders" and uses them to find the market-clearing prices. This avoids the difficulty of actually specifying the actual trading mechanism but hides the effect of the market microstructure on the process of price formation. In the next lecture, we will see that the microstructure literature improves on the rational expectations model by making explicit the process of price determination.

All in all, models based on rational expectations are difficult to construct and difficult to interpret. The common approach circumvents these difficulties by using specific examples. While this makes things tractable, the approach is special in the sense that even smaller deviations from the assumptions in the example may change the equilibrium drastically, or make it disappear.

2. Market microstructure: Kyle's model

2.1 Introduction

The models on market microstructure differ on the assumptions made on how the best available price is set. In auction models (also known as "order-driven"), the best available price is defined by the submitted orders. In dealer models (also known as "quote-driven"), it is defined by dealer quotes.

The rational expectations model can be seen as an example of an (implicit) auction model. Kyle's (1985) model improves on this by making it explicit the process by which prices are formed, while keeping the auction structure and the use of an expectations consistent pricing rule. In particular, Kyle assume batch-clearing; that is, all orders are fulfilled simultaneously at the same price.

The model assumes that there is a *market-maker* (named Secunda) who set prices and thus acts as an auctioneer. Moreover, the market-maker can take trading positions and has privileged access to information on the order flow. This changes the nature of the pricing rule because the act of price setting is assigned to a player within the model. The market-maker must set prices using only the information which is available to him, which is determined by the trading protocol. This generates a relationship between the price and the trading protocol.

Besides the market-maker, Kyle assumes that there is one informed agent (named Primus) and a number of liquidity traders in the market. The market maker aggregates the orders and clears all trades at a single price. The informed trader chooses those transactions which maximize the value of his private information. This provides a relationship between the price and the strategic use of information by the informed trader. Thus price reflects both the trading protocol and the strategic behavior of the informed trader.

2.2 The model

We consider a one-asset one-period economy, with a zero riskless interest rate. There are three types of agents: one informed trader (Primus), one uninformed market-maker (Secunda), and many noise traders who trade only for liquidity or hedging reasons. Primus and Secunda are risk-neutral expected utility maximizers. All random variables are independent and normally distributed, with strictly positive standard deviations. The only available asset is a risky stock which will pay a risky amount $Y \sim N(m, s_y)$ at the end of the period.

Primus is the informed trader, who receives a perfectly precise signal about Y and learns that $Y = y$. Secunda is the uninformed market-maker, who knows only the prior distribution of Y. After Primus learns that $Y = y$, market orders from Primus and the noise traders are submitted to Secunda to be executed at a uniform market-clearing price p. The noise traders submit a random demand $Q^u \sim N(0, s_u)$; if this is negative, they are on balance selling. Primus submits a demand Q^i without observing the realization of Q^u. The market-maker receives an aggregate demand $Q = Q^i + Q^u$; she knows the sum but not who demanded what.

The market-maker's pricing rule mandates that she earns zero profits. This is consistent with free entry of competing market-makers, which impairs any monopoly power of the single market-maker. This implies that the market-maker sets prices such that $p = E(Y|Q)$.

Primus places an order Q^i which maximizes his profit $E[(Y-p)Q^i|Y=y] = (y-p)Q^i$. Primus' order is influenced by the price quoted by the market maker. At the same time, his demand Q^i affects the price quoted. This strategic interaction between the market-maker and the informed trader is what makes the model tick: Secunda's choice of p depends on Q^i and Primus' choice of Q^i depend on p.

Kyle proves that there exists a linear equilibrium for this model such that the market-maker pricing rule is

$$P(Q) = m + \alpha Q \tag{10}$$

and Primus' trading rule is

$$Q^i = \beta(y - m), \tag{11}$$

where

$$\alpha = \frac{1}{2}\frac{s_y}{s_u} \qquad \text{and} \qquad \beta = \frac{s_u}{s_y}. \tag{12}$$

We will show momentarily a piece of the argument which establishes this result. Before that, a quick commentary may be useful.

Note that both the pricing and the trading rule depend on the same parameters (although their ratio is inverted). When α is high and orders have a significant price impact, then β is low because Primus trades less aggressively (to avoid the impact of his own trades). When s_y is high, Primus' information is more likely to be substantial and therefore Secunda adjusts price more aggressively. When s_u is high, Primus' order is less likely to be a conspicuous component of the total order flow, and therefore he can afford to trade more aggressively.

To show that (10) and (11) constitute an equilibrium, we need to prove two facts. First, that the best reply of Secunda to an insider using the trading rule (11) is precisely (10). Second, that the best reply of Primus to a market-maker using the pricing rule (10) is precisely (11). We prove only the first fact, and leave the second as an exercise.

Thus, assume (11). Secunda's prior for Y is that $Y \sim N(m, s_y)$. Since she observes the aggregate demand $Q = Q^i + Q^u$, she can use this information to update her prior. As $Q^i = \beta(y-m)$, the total order flow can be written $Q = \beta(y-m) + Q^u$, with $Q^u \sim N(0, s_u)$. Dividing Q by β (which is known in equilibrium) and adding m (which is known), Secunda can construct the observable random variable

$$\eta := m + \frac{Q^i + Q^u}{\beta} = y + \frac{Q^u}{\beta},$$

which is normally distributed with mean y and standard deviation $s_\eta = s_u/\beta = s_y$. This can be used to make inferences about Y. Exploiting known results from Lecture 7 (with precisions instead of variances), we know that

$$Y|\eta \sim N\left(\frac{m\tau_y + \eta\tau_\eta}{\tau_y + \tau_\eta}, \frac{1}{\tau_y + \tau_\eta} \right).$$

Since $\tau_\eta = \tau_y$, we obtain

$$E\,(\,Y\,|\,\eta\,) = \frac{m\tau_y + \eta\tau_\eta}{\tau_y + \tau_\eta} = \frac{m + \eta}{2}.$$

Secunda sets her price equal to her best estimate of Y; that is, $P = E(Y|\eta)$. By definition, $\eta = m + (Q/\beta)$. Hence, the price set by Secunda is:

$$P = E\,(\,Y\,|\,\eta\,) = \frac{1}{2}\left(m + \frac{Q}{\beta} + m\right) = m + \frac{1}{2\beta}Q.$$

Using again the value of β from (12), it is easy to check that this pricing rule indeed matches (10), as it was to be shown.

2.3 Lessons learned

There are a few major insights to be gained from Kyle's model.

1. When setting the price, the market-maker explicitly updates the analysis of fundamentals (corresponding to her prior) with the information embedded in the order flow. She uses information gathered by the traders as transmitted by their demands. Order flow communicates information about fundamentals because it contains the trade of those who analyze/observe fundamentals.

We can actually measure how much of the trader's information is revealed in Kyle's model by looking at the variance of Secunda's posterior distribution for Y. Before she observes the order flow, this variance is s_y. After she observes Q and updates her prior, the variance becomes

$$V(Y|\eta) = \frac{1}{\tau_y + \tau_\eta} = \frac{s_y^2}{2}.$$

This is exactly half of the prior variance s_y^2. Regardless of the exact value, the important message is that the updated variance is somewhere in between the prior variance and a zero variance. If the updated variance were to remain s_y^2, Secunda would learn nothing and Primus could make infinite profits. If it were to drop to zero, Primus would make no profit because all his trades would clear at the perfectly revealing price Y.

2. Since the market-maker cannot separate informative from uninformative trade, the transmission of information is noisy and the informed trader can use this to his advantage. Primus can partially hide his trade from the market-maker in the order flow from the noise traders, reducing the amount by which price moves against him.

3. Liquidity and market efficiency are deeply related. Efficient markets tend to gravitate towards constant liquidity, defined as the *price impact* of orders. To see why, suppose that liquidity is not constant: for example, suppose it is known that the market will be fairly illiquid in the next month and then will revert to standard liquidity. Then if Primus buys 10's worth of stock each day in the next month, each purchase will push the price of the stock progressively upward. This is because, if Primus' trades communicate information about fundamentals, these price increases induced by the order flow should persist. Then, when at the end of next month liquidity has returned to normality, Primus could suddenly sell the 300's worth of stock purchased during the month with almost no price impact and make a riskless excessive profit similar to an arbitrage opportunity. An efficient market should prevent making excessive profits.

Kyle (1985) considers a multiple-period extension of his model which generates constant liquidity in equilibrium. Intuitively, this follows because the informed trader must balance the effects of his current trades on his future trading opportunities: if he trades too much too soon, the price will adjust rapidly and his profits will be smaller.

3. Market microstructure: Glosten and Milgrom's model

3.1 Introduction

The market in Kyle's model is "order-driven": the market-maker sees the order flow and sets a price which clears the market in a single batch. Demand and supply meet at a single price and simultaneously. This lecture considers a different trading protocol, possibly closer to reality, which is "quote-driven" and involves sequential trades. Unlike Kyle's auction, this sequential-trade model describes a dealership market.

All trades involve a dealer, who posts bid and ask prices. Traders arrive sequentially and can trade at the current bid-ask prices. Thus orders are fulfilled sequentially at (possibly) different prices. After each trade, the dealer updates his bid and ask prices to reflect the information he has learned from his privileged position. At each trade, the current trader may be informed or uninformed so the model can accomodate more than one informed trader.

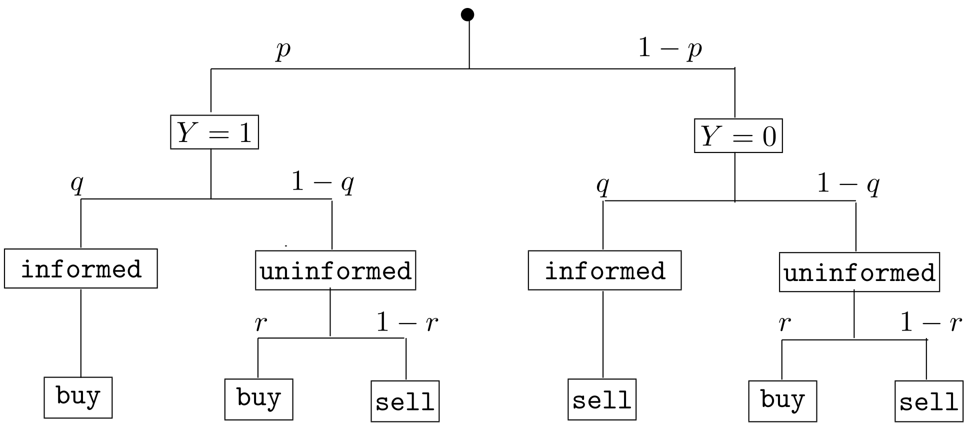
This trading protocol implies some important differences from Kyle's model. First, there is an explicit bid-ask spread, as opposed to the single market-clearing price of Kyle's model. Second, the spread occurs even if the dealer is risk-neutral and behaves optimally incorporating in the price all the information he can extract from the order flow. Third, since trades occur sequentially, it is possible to analyze explicitly how the information content of each trade affects the bid-ask spread.

3.2 The model

We consider a one-asset one-period economy, with a zero riskless interest rate. There are three types of agents: a single dealer (Primus), several informed traders and many uninformed traders. Both the dealer and the informed traders are risk-neutral expected utility maximizers, while the uninformed traders trade only for liquidity or hedging reasons.

The only available asset is a risky stock which will pay a risky amount Y at the end of the period. For simplicity, we assume that it can take only the high value $Y = 1$ or the low value $Y = 0$. The prior distribution is $P(Y = 1) = p$ and $P(Y = 0) = 1 - p$. All the informed traders know the realization of Y (because they have received a perfectly informative signal about it), while Primus does not.

Trading is organized as a sequence of bilateral trading opportunies, taking places one after the other but always within the end of the period. The pool of potential traders is given and the dealer knows that q of it are informed traders and $(1 - q)$ are uninformed traders. At each trading opportunity, one trader is randomly chosen from this pool (with replacement) and is offered the chance to buy or sell one unit of the stock at the current bid or ask price. The informed trader can buy, sell, or pass at her discretion. As for the uninformed traders, we assume for simplicity that they buy with probability r or sell with probability $1 - r$. The figure gives a pictorial representation of the process (under the assumption that an informed trader buys when he knows that $Y = 1$ and sells otherwise).



As in Kyle's model, the dealer sets prices such that the expected profit on any trade is zero. (This can be justified by assuming a competitive dealers' market.) This implies that Primus must set prices equal to his conditional expectation of the asset's value given the type of transaction taking place.

More specifically, before trading opportunity t occurs, Primus must offer a bid price b_t and an ask price a_t such that

$$b_t = E(Y \mid \text{next trader sells}) \qquad \text{and} \qquad a_t = E(Y \mid \text{next trader buys}). \tag{13}$$

This rule takes explicitly into account the effect that the sale/purchase of one unit would have on Primus' expectations. This makes sure that his prices are "regret-free", in the sense that — given the trade that actually occurs — the dealer believes that the price is fair.

Such "regret-free" price-setting behavior makes sure that prices incorporate the information revealed by a trade. Due to the signal value of each trade, as trading goes on, the dealer keeps revising his beliefs and sets new trading prices. This generates a sequence of bid-ask prices $\{b_t, a_t\}$ that change over time, paralleling the evolution of Primus' beliefs. Let us work out an example and see what happens.

3.3 An example

Suppose $p = q = r = 1/2$. Denote by B_t and S_t respectively the event that at the trading opportunity t there is a buy or a sale. By (13), the ask price at the first trading opportunity should be

$$a_1 = E(Y \mid B_1) = 1 \cdot P(Y = 1 \mid B_1) + 0 \cdot P(Y = 0 \mid B_1), \tag{14}$$

while the bid price should be

$$b_1 = E(Y \mid S_1) = 1 \cdot P(Y = 1 \mid S_1) + 0 \cdot P(Y = 0 \mid S_1). \tag{15}$$

In order to find what a_1 should be, we need to compute $P(Y = 1 \mid B_1)$. By Bayes' rule,

$$P(Y = 1 \mid B_1) = \frac{P(Y = 1) \cdot P(B_1 \mid Y = 1)}{P(Y = 1) \cdot P(B_1 \mid Y = 1) + P(Y = 0) \cdot P(B_1 \mid Y = 0)}. \tag{16}$$

Since we know by assumption that $P(Y = 1) = P(Y = 0) = 1/2$, it suffices to determine $P(B_1 \mid Y = 1)$ and $P(B_1 \mid Y = 0)$. We know that the uninformed traders buy always with probability 1/2. On the other hand, the informed traders know Y and therefore buy only if Y is high and sell only if Y is low. That is, they buy with probability 1 if $Y = 1$ and sell with probability 1 if $Y = 0$.

Therefore, conditional on $Y = 1$, the probability of a buy is 1/2 if it comes from an uninformed trader and 1 if it comes from an informed trader. Since uninformed and informed traders are equally likely to come up for trade, the overall probability is $P(B_1 \mid Y = 1) = (1/2)\cdot(1/2)+(1/2)\cdot 1 = 3/4$. By a similar reasoning, $P(B_1 \mid Y = 0) = (1/2)\cdot(1/2)+(1/2)\cdot 0 = 1/4$. Substituting in (16), we find

$$P(Y = 1 \mid B_1) = \frac{(1/2) \cdot (3/4)}{(1/2) \cdot (3/4) + (1/2) \cdot (1/4)} = \frac{3}{4}.$$

By a similar reasoning, one can deduce that $P(Y = 1 \mid S_1) = 1/4$.

By (14) and (15), we obtain that the dealers sets ask and bid prices respectively equal to

$$a_1 = \frac{3}{4} \qquad \text{and} \qquad b_1 = \frac{1}{4}.$$

Suppose that there actually occurs a buy at a price of 3/4. What will be the new bid and ask prices? The probability distribution for Y after a buy is $P(Y = 1 \mid B_1) = 3/4$ and $P(Y = 0 \mid B_1) = 1/4$. This acts as a new prior for the next trading opportunity. Then

$$a_2 = E(Y \mid B_1, B_2) = 1 \cdot P(Y = 1 \mid B_1, B_2) + 0 \cdot P(Y = 0 \mid B_1, B_2)$$

and

$$b_2 = E(Y \mid B_1, S_2) = 1 \cdot P(Y = 1 \mid B_1, S_2) + 0 \cdot P(Y = 0 \mid B_1, S_2).$$

We need to compute $P(Y = 1 \mid B_1, B_2)$ and $P(Y = 1 \mid B_1, S_2)$. By Bayes' rule,

$$P(Y = 1 \mid B_1, B_2) = \frac{(3/4) \cdot (3/4)}{(3/4) \cdot (3/4) + (1/4) \cdot (1/4)} = \frac{9}{10},$$

which in turn implies $a_2 = 9/10$.

Note that, if a sale had occurred instead of a buy, the price would have been set to $a_2 = 1/2$ and $b_2 = 1/10$. Therefore, the fact that the first transaction is a buy or a sale reveals information. On the other hand, since $(a_t - 1)(b_t - 0) \neq 0$ for all t, information is never fully revealed.

The exercise can be repeated. The exact sequence of bid-ask prices will depend on the actual trading events. For instance, if the first four events are B_1, S_2, B_3, and B_4, the sequence will be

t	b_t	a_t	trade
1	1/4	3/4	B_1
2	1/2	9/10	S_2
3	1/4	3/4	B_3
4	1/2	9/10	B_4
5	3/4	27/28	etc.

A sufficient statistics for the current bid and ask prices is the difference between the number β_t of buys occurred before t and the number σ_t of sales occurred before t. For instance, if $\beta_t - \sigma_t = 0$, then $b_t = 1/4$ and $a_t = 3/4$.

3.4 Comments on the model

The assumptions about the trading protocol are crucial. Informed traders profit from trading if prices do not yet reflect all the available information. An informed trader prefers to trade as much (and as often) as possibile. By so doing, informed traders quickly reveal information which is immediately incorporated in prices.

This cannot occur in the model because the only trader who is allowed to trade is chosen randomly and she can only buy or sell one unit of stock. Thus, if an informed trader desires to trade further, she must return to the pool of traders and wait to be selected again.

The probabilistic selection process dictates that the population of traders facing the dealer is always the same as the population of potential traders. This makes it possible for the dealer to know the probability that he is trading with an informed trader. Moreover, it implies that plausible trading scenarios are ruled out. For instance, whenever information is likely to become more dispersed over time, the fraction of informed traders should increase with time (and the dealer would need to learn another parameter yet). This cannot occur here.

An important result of the model is that transaction prices form a martingale. That is, the best predictor for the transaction price in $t+1$ (given the information I_t available after trade t) is the transaction price in t: $E(p_{t+1}|I_t) = p_t$.

For instance, in the example above, consider the situation after having observed the first buy. The current price is $p_t = a_t = 3/4$. The next transaction may come from an uninformed trader with probability 1/2 or from an informed trader with probability 1/2. The probability that it will be a buy is 1/2 if the next trader is uninformed and 3/4 if he is informed. Therefore, the probability that the next transaction will be a buy is $(1/2) \cdot (1/2) + (1/2)(3/4) = 5/8$. Hence, the next price will be 9/10 with probability 5/8 and 1/2 with probability 3/8. Then $E(p_{t+1}|I_t) = (9/10) \cdot (5/8) + (1/2) \cdot (3/8) = 3/4 = p_t$.

The martingale property dictates that prices respect semi-strong efficiency, in the sense that they reflect all the information available to the dealer. It can be shown that, in the limit, all information is revealed and $a_t - b_t \to 0$ with both a_t and b_t converging to 0 or 1 depending on whether, respectively, $Y = 0$ or $Y = 1$.

3.5 Lessons learned

There are a few major insights to be gained from Glosten and Milgrom's model.

1. Information alone is sufficient to induce spreads, independently of the risk attitude of the dealer or of his inventory costs. The equilibrium spread in this model is such that when the dealer trades with an informed trader he loses money (as in Kyle's model, the informed trader knows exactly the value of the asset). To prevent overall losses, the dealer must offset them with gains from trading with uninformed traders. The equilibrium spread balances these losses and gains exactly so that expected profits are zero.

2. Learning takes place over time, as it involves the sequential arrival of distinct orders. The dealer does not know whether behind a single trade there is an informed trader who knows something that the dealer does not, or an uninformed trader who must trade for reasons unrelated to fundamentals. However, if a preponderance of sales takes place over time, the dealer adjusts her beliefs and prices downward. The private information gradually finds its way in the dealer's prices.

3. There is a process of price discovery behind the (dynamic) adjustment to market efficiency. The dealer must uncover the private information hidden behind trades before market prices can be efficient.

References

[1] F. de Jong and B. Rindi (20009), *The Microstructure of Financial Markets*, Cambridge University Press.

[2] L. Glosten and P. Milgrom (1985), "Bid, ask, and transaction prices in a specialist market with heterogenously informed traders," *Journal of Financial Economics* **13**, 71–100.

[3] S.J. Grossman and J.E. Stiglitz (1980), "On the impossibility of informationally efficient markets," *American Economic Review* **70**, 393–408.

[4] A. Kyle (1985), "Continuous auctions and insider trading," *Econometrica* **53**, 1315–1335.

[5] R.K. Lyons (2001), *The microstructure approach to exchange rates*, The MIT Press, Chap. 4.

[6] M. O'Hara (1995), *Market microstructure theory*, Basil Blackwell.