Lecture notes

Pierre Collin-Dufresne

Swiss Finance Institute

EPFL

pierre.collin-dufresne@epfl.ch

October 28, 2020

1 Introduction

These lectures will cover classic papers that study the role of information frictions and illiquidity for asset returns and volume. Financial markets are characterized by large daily trading volumes and episodic market crashes and 'bubbles.' Stories abound about successful traders who make fortunes by pursuing 'clever' trading strategies. At the same time, there is the occasional bankruptcy or ponzi scheme to suggest it may actually be difficult to 'beat the market.' It is natural to investigate whether such evidence is consistent with models of differentially informed investors. We will start with model where agents share common priors but receive different signals and then investigate models where agents have different priors and thus different posterior beliefs even when they receive the same signal.

2 Efficient markets, information, and the role of financial prices

Hayek (1945) points to the role of the price system to aggregate dispersed information. Information about 'circumstances of time and place' that are difficult to summarize with statistics and to convey to a central planner. Argues that the price system uniquely can make agents in dispersed location and at different times coordinate their actions and optimize the 'utilization of knowledge not given to anyone in totality". Importantly, " it does not matter why the price of tin is high" for agents to start to save on tin and substitute for other inputs.

Related, Fama's efficient market hypothesis (1976) contends that market prices always reflect all available information efficiently, thus leading to optimal resource allocation. The notion of market efficiency goes back at least to Bachelier (1900) and the idea that trading by agents leads prices to reflect all available information. It usually leads to some assumption about the 'unpredictability' of price movements (the 'martingale' condition), based on different levels of information (prices only, public, private).

3 No-trade theorems

This section is based on the papers by Tirole (1982) and Milgrom and Stokey (1982). We establish that in a market where investors share common priors and trade solely for speculative motives, there can be no-trade in equilibrium even if investors have different informative signals.

3.1 No trade and pure speculation

Suppose n agents with (random) terminal endowment W_i can trade and asset at price P_0 with terminal payoff P_1 which is uncorrelated with W_i $\forall i$. Assume that all agents are strictly risk-averse ($U' > 0, U'' < 0$) and $\max_x E[U(W_i + x_i(P_1 - P_0(1+r_f))) \,|\, \phi_i, \Phi]$, where ϕ_i is a private signal observed by agent i and Φ is all publicly available information (which includes the trading price P_0).

The FOC for agent i gives :

$$E[U'(W_i + x(P_1 - P_0(1 + r_f)))(P_1 - P_0(1 + r_f)) \,|\, \phi_i, \Phi] = 0$$

Since agents are always free to not trade it must be that:

$$x_i > 0 \iff E[P_1 \,|\, \phi_i, \Phi] > P_0(1 + r_f)$$

Market clearing implies $\sum_i x_i = 0$. For simplicity let's assume there are only $n = 2$ traders. Then suppose there is trade and $x_1 = -x_2 > 0$ (say). It thus follows that

$$E[P_1 \,|\, \phi_1, \Phi] > P_0(1 + r_f) > E[P_1 \,|\, \phi_2, \Phi]$$

and by the law of iterated expectation we obtain the contradiction:

$$E[P_1 \,|\, \Phi] > E[P_1 \,|\, \Phi]$$

Therefore any rational expectation equilibrium must result in *no trade* and furthermore

$$P_0(1 + r_f) = E[P_1 \,|\, \phi_i, \Phi] = E[P_1 \,|\, \Phi] \quad \forall i$$

This no-trade theorem shows that if it is common-knowledge that all agents starting from common priors want to trade for purely speculative reasons based on private signals, then in equilibrium they should all refrain from trading and prices should adjust to incorporate their various signals so that no-trade is an equilibrium. Notice that no-trade does not imply no-price change!

Where does the argument fail if the end of period price P_1 is correlated with the agents terminal wealth W_i?

Where does the common-knowledge assumption play a role in the argument?

For trade to take place in equilibrium, agents must differ in terms of endowments, preferences, or beliefs.

3.2 No-trade and pareto efficient equilibrium: Milgrom and Stokey (1982)

Consider an economy with two dates: agents consume at date 0 and date 1 in states $1, \ldots, S$. They have endowments in all states. Markets are effectively complete, so that at date 0 there is a round of trading whereby agents achieve a pareto optimal allocation and consume their date 0 consumption. After that, but before date 1, private information ϕ_i arrives to all traders (but there is no change to the state space and initial endowments) and markets reopen for trade. Agents will thus condition their trades on their private information ϕ_i and all new publicly availalbe information Φ (which includes prices). We use the following notation:

- W_i is agent i's optimal date 1 total consumption after the first round of trading. It's a random variable (agent i plans to consume W_{is} in state s).

- $p(s, \phi_i, \Phi)$ is the probability that state s occurs and agent i gets information ϕ_i and public information is Φ. Importantly, all agents have common prior. So they all agree on these probabilities. They may not know what information agent i actually obtains, but ex ante they know every other agents possible information set and they agree about the prior probabilities.

- $q(s)$ is the ex-ante state price of state s and $\hat{q}(s)$ is the new price system that prevails after the information is released.

Theorem 2 of MS states that there will be no trading when markets reopen.[1] That is, in spite of the new information, the initial allocation is still Pareto optimal and there are no available gains from trading.

Indeed, assume that there is a globally feasible trade $t_i(\phi_i, \Phi)$ (a random variable where agent i trades t_{is} in states s conditional on her private information) , so that

$$\sum_i t_i(\phi_i, \Phi) = 0 \tag{1}$$

$$E[U_i(W_i + t_i)|\phi_i, \Phi] \geq E[U_i(W_i|\phi_i, \Phi)] \tag{2}$$

Then consider another trade $T_i = \sum_{\phi_i,\Phi} p(\phi_i, \Phi) t_i(\phi_i, \Phi)$ (which is essentially the ex-ante expectation of the assumed superior ex-post trades). Clearly T_i does not depend on future private of public information. Further, $\sum_i T_i = \sum_i \sum_{\phi_i,\Phi} p(\phi_i, \Phi) t_i(\phi_i, \Phi) = \sum_{\phi_i,\Phi} p(\phi_i, \Phi) \sum_i t_i(\phi_i, \Phi) = 0$. So clearly the trade is globally feasible. Further

$$U_i(W_{is} + T_{is}) = U_i(W_{is} + \sum_{\phi_i,\Phi} p(\phi_i, \Phi) t_{i,s}(\phi_i, \Phi)) \geq \sum_{\phi_i,\Phi} p(\phi_i, \Phi) U_i(W_{is} + t_{i,s}(\phi_i, \Phi))$$

By Jensen's inequality and given the concavity of the utility function. Then taking unconditional expectation we get

$$\begin{aligned} E[U_i(W_i + T_i)] \geq \sum_s p(s) \sum_{\phi_i,\Phi} p(\phi_i, \Phi) U_i(W_{is} + t_{i,s}(\phi_i, \Phi)) &= \sum_{\phi_i,\Phi} p(\phi_i, \Phi) \sum_s p_s U_i(W_{is} + t_{i,s}(\phi_i, \Phi)) \\ &= E[E[U_i(W_i + t_i)|\phi_i, \Phi)]] \\ &= E[U_i(W_i + t_i)] \geq E[U_i(W_i)] \end{aligned}$$

[1] Here we follow the proof given in Ingersoll (1987)

(where the last line follows from taking the unconditional expectation in inequality 2 above) which contradicts the Pareto optimality of the initial trade t_i (note that if $T_i > 0$ then by strict concavity of the utility function $T_i/2$ will be strictly Pareto superior to t_i).

Note that the fact that agents don't want to retrade when market reopen (as long as endowments, preferences, or states are not changed by the release of new information) does not mean that prices don't change. In fact, state prices change so that they reveal all the private information of investors, in the sense that conditionning on the common public price gives each agent as much information as she could get based on her private signal. In that sense the new equilbrium price is 'fully revealing.' More specifically, we will show that

$$p(s|\phi_i, \Phi) = p(s|\hat{q})$$

where $\hat{q}_s$ are the new state prices. To see this note that **ex-ante** Pareto optimality implies

$$p(s)U'_{is}(W_{is}) = \lambda_i q_s \;\; \forall i, s$$

Ex-post Pareto-optimality of the **same** allocation implies

$$p(s|\phi_i, \Phi)U'_{is}(W_{is}) = \hat{\lambda}_i \hat{q}_s \;\; \forall i, s$$

Thus

$$p(s|\phi_i, \Phi) = \frac{\hat{\lambda}_i \hat{q}_s}{\lambda_i q_s} p(s) \;\; \forall i, s$$

and

$$\frac{p(s|\phi_i, \Phi)}{p(s'|\phi_i \Phi)} = \frac{\hat{q}_s q_{s'} p(s)}{\hat{q}_{s'} q_s p(s')} \;\; \forall s, s'$$

note that the RHS is independent of i and depends only on ex-ante prior information and new prices. Thus each agent can fully determine the posterior probabilities $p(s|\phi_i, \Phi)$ solely from the initial common prior state probabilities and common-knowledge new price system $\hat{q}$.

The conclusion is that trading on private information, based on purely speculative motives (i.e., starting from a Pareto optimal allocation and not changing preferences, endowments or states), will not occur in a rational expectations' equilibrium. However, (state-)prices will adjust. In fact they will "swamp" the private signals received by any agent, in that, after equilibrium prices are determined each agent can afford to forget the signal he observrved to compute his posterior beliefs. she only needs to know the change in prices.

MS conclude:

> *Our results concerning rational expectations market equilibria raise the disturbing questions expressed by Beja , Grossman and Stiglitz and Tirole: Why do traders bother to gather information if they cannot profit from it? How does information come to be reflected in prices if informed traders do not trade or if they ignore their private information in making inferences? These questions can be answered satisfactorily only in the context of models of the price formation process, our central result, the no-trade theorem, applies to all such models when rational expectations are assumed.*

4 The Grossman REE paradox (Grossman (1976))

- N agents with wealth $W_{i0} = X_{iF} + X_i P_0$ can invest in risk-free asset with gross return R and risky asset with price P_0 which pays off $P_1 \sim N(\overline{P}, \sigma_p)$ at time 1.

- $W_{i1} = RW_{i0} + (P_1 - RP_0)X_i$

- Each agent i observes signal $y_i = P_1 + \epsilon_i$. $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ iid with $Cov(\epsilon_i, \epsilon_j) = 0$.

- Assume agents are CARA normal with risk-aversion a_i. Since terminal wealth will be normally distributed in equilibrium, agents have mean-variance preferences.

$$\max_X E[W_{i1} \,|\, \mathcal{I}_i] - \frac{a_i}{2} V[W_{i1} \,|\, \mathcal{I}_i] = \max_X RW_{i0} + X_i(E[P_1 \,|\, \mathcal{I}_i] - RP_0) - \frac{a_i}{2} X_i^2 V[P_1 \,|\, \mathcal{I}_i]$$

- The FOC gives

$$X_i = \frac{E[P_1 \,|\, \mathcal{I}_i] - RP_0}{a_i V[P_1 \,|\, \mathcal{I}_i]}$$

- An equilibrium is a price $P(s_1, \ldots, s_n)$ such that markets clear $\sum_i X_i = \overline{x}$, where $\overline{x}$ is the total supply of the risky asset.

- Grossman discusses how the equilibrium price functional would arise from a "tatonnement" process of a sequence of approximations to the equilibrium price (see page 577).

- To solve for the equilibrium, we will follow the three step procedure. First, conjecture an equilibrium price functional. Second, given this conjecture, solve the first order condition of each agent who conditions her demand on the information contained in price. Third, clear markets and verify that the market clearling price is as conjectured in the first step.

- Guess an equilibrium price functional $P_0(y) = \alpha_0 + \sum_i \alpha_i y_i$. Now, because of the symmetry it is natural to assume that α_i are identical and to posit the simpler form $P_0(y) = \alpha_0 + \alpha\overline{y}$, where $\overline{y} = \frac{1}{n}\sum_i y_i$. Note that $\overline{y} = P_1 + \overline{\epsilon}$, where $\overline{\epsilon} \sim N(0, \frac{\sigma_\epsilon^2}{n})$.

- Given our conjecture the information set of each agent is $\mathcal{I}_i = \{y_i, P_0(y)\} = \{y_i, \overline{y}\}$. Our first step is to show that:

$$E[P_1 \,|\, \mathcal{I}_i] = E[P_1 \,|\, y_i, \overline{y}] = E[P_1 \,|\, \overline{y}]$$

To prove this we will use the standard Gaussian projection theorem (see subsection 4.1). P_1 can be decomposed:

$$P_1 = \overline{P} + \beta(\overline{y} - E[\overline{y}]) + \nu$$

where $\nu \perp \overline{y}$ (that is $E[\nu\overline{y}] = 0$). The 'regression' coefficient is given by:

$$\beta = \frac{Cov(P_1, \overline{y})}{V[\overline{y}]} = \frac{\sigma_P^2}{\sigma_P^2 + \frac{\sigma_\epsilon^2}{n}}$$

Note also that $E[\overline{y}] = \overline{P}$. By orthogonality we have

$$V[\nu] = V[P_1] - \beta^2 V[\overline{y}] = \frac{\sigma_P^2 \frac{\sigma_\epsilon^2}{n}}{\sigma_P^2 + \frac{\sigma_\epsilon^2}{n}} = \frac{1}{\tau_p + n\tau_\epsilon}$$

where we have used the notation that the precision of a random variable $\tau_\nu = \frac{1}{\sigma_\nu^2}$.

Now, clearly $E[P_1|\overline{y}] = \overline{P} + \beta(\overline{y} - \overline{P})$. We want to show that $E[P_1|\overline{y}, y_i] = \overline{P} + \beta(\overline{y} - \overline{P})$ as well. For that it is sufficient to show that $E[\nu \,|\, y_i] = 0$. And this follows from

$$E[(P_1 - \overline{P} - \beta(P_1 - \overline{P} + \overline{\epsilon}))(P_1 + \epsilon_i)] = \sigma_P^2(1-\beta) - \beta\frac{\sigma_\epsilon^2}{n} = 0$$

- It follows then that

$$E[P_1|\mathcal{I}_i] = \overline{P} + \beta(\overline{y} - \overline{P})$$
$$V[P_1|\mathcal{I}_i] = \sigma_\nu^2 = \frac{1}{\tau_p + n\tau_\epsilon}$$

So we get the demand of each agent (given our conjectured price functional)

$$X_i = \frac{\overline{P} + \beta(\overline{y} - \overline{P}) - RP_0}{a_i \sigma_\nu^2}$$

And market clearing $\sum_i X_i = \overline{x}$ gives

$$\overline{P} + \beta(\overline{y} - \overline{P}) - RP_0 = \bar{a}\sigma_\nu^2 \overline{x}$$

where $\bar{a} = \frac{1}{\sum_i \frac{1}{a_i}}$ is the harmonic average of the risk-aversion coefficient (the 'representative' CARA agent's risk-aversion).

- Note the remarkable feature that plugging back the market clearing relation into the demand we see that each individual demand becomes $X_i = \frac{\bar{a}\overline{x}}{a_i}$, which does not actually depend on price! This is because of the *information effect* noted by (Admati (1989)). When prices increase (say) in most models there is a wealth (absent here due to CARA) and substitution effect. Here in addition there is an information effect which exactly offsets the substitution effect (higher prices mean better (information about) fundamentals).

- Note the Skizophrenia of the investors. They take prices as given, i.e., ignore the impact of their demand on price. On the other hand, they use prices to learn about other traders signals. But prices can only reflect other traders information if these have an impact on prices. See Hellwig (1980) who tries to rationalize this by considering an economy with infinitely many small traders.

- We then get the equilibrium price functional consistent with optimality and market clearing;

$$P_0(\overline{y}) = \frac{1}{R}\left(\overline{P}(1-\beta) + \beta\overline{y} - \bar{a}\sigma_\nu^2\overline{x}\right)$$

We verify that it is indeed consistent with our initial conjecture and have thus found an equilibrium. Is it unique? (yes among 'linear equilibria', but there might exist non-linear equilibria where the price is not a linear function of all the signals... See DeMarzo and Skiadas (1998) and also "Multiple Equilibria in Noisy Rational Expectations Economies" by Domotor Palvolgyi and Gyuri Venter (2015)).

- Grossman points out the paradox of the equilibrium. All agents learn better information (a sufficient statistic which dominates any single piece of dispersed information) from prices and can thus discard their own information. But if they did so, then how would information get into price in the first place? The issue is that all traders behave competitively and ignore the impact they have on price (through their trading based on their own information). they take the price functional 'as given'.

- Theorem 2: suppose $P^*(\overline{y})$ is linear equilibrium function and $P^{**}(\overline{y})$ another increasing equilibrium functional then there must exist $H(\dot{)}$ increasing so that $P^{**} = H(P^*)$. If H is not identity function then P^{**} cannot be an equilibrium. This follows directly from the market clearing condition using the FOC:

$$\overline{x} = \sum_i \frac{E^i[P_1] - RP^*}{aV[P_1]} < \sum_i \frac{E^i[P_1] - RP^{**}}{aV[P_1]}$$

  If the LHS clears markets then the RHS cannot. The implication is that two equilibrium price functionals cannot give access to the same information $\overline{y}$ for then the expectations and variances are identical and the LHS and RHS will then only differ by the price component, which cannot differ. So for there to be different equilibria, the price functional has to disclose different information to agents.

- Grossman also discusses the Pareto efficiency of the equilibrium. He points out that the decentralized equilibrium gives each investor access to $\overline{y}$ a sufficient statistic, which dominates all the individual information signals. So a central planner having access to all y_i would not do any better.

- Grossman concludes with a critique of Hayek (1945). Unlike Hayek, who argues that agents shouldn't care why prices are high as long it makes them reduce their consumption of a more expensive good, Grossman points out that without noise equlbrium would typically break down if informatino acquisition is costly, and with noise, then prices would not be fully revealing and therefore agents would care 'why' prices are high (whether for fundamental or noise reasons).

4.1 The Gaussian Projection theorem

- The multi-variate Gaussian projection theorem: Suppose $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is a multivariate normal vector of random variables (X_1 is N-dimensional and X_2 is K-dimensional).
- The expected return vector is $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$.
- We partition the covariance matrix $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$
- Then the vector of random variable X_1 conditional on $X_2 = q$ is Gaussian $(X_1 \mid X_2 = q) \sim N(\bar{\mu}, \bar{\Omega})$ with

$$\begin{aligned} \bar{\mu} &= \mu_1 + \Omega_{12}\Omega_{22}^{-1}(q - \mu_2) \\ \bar{\Omega} &= \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21} \end{aligned}$$

- This is similar to multi-variate linear regression. The proof follows.
- Define $X_1^e = X_1 - \mu_1$ and $X_2^e = X_2 - \mu_2$.
- Note that we can decompose $X_1^e = BX_2^e + u$ where we choose the (N, K) matrix B so that u is orthogonal to any element of X_2, that is: $E[(X_1^e - BX_2^e)X_2^\top] = \mathbf{0}_{(N,K)}$. Equivalently $BE[X_2^e X_2^\top] = E[X_1^e X_2^\top]$, and thus $B = \Omega_{12}\Omega_{22}^{-1}$.
- Since X_1, X_2 are jointly normal, it follows that u is normally distributed with zero mean and with variance $Var(u) = Var(X_1) - Var(BX_2) = \Omega_{11} - B\Omega_{22}B^\top = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$.
- It follows from the orthogonal decomposition above that the distribution of X_1 conditional on $X_2 = q$ is normal with mean $E[X_1 \mid X_2 = q] = \mu_1 + B(q - \mu_2)$ and with variance $Var[X_1 \mid X_2 =$

$q] = Var[u]$

5 On the impossibility of informationally efficient markets (Grossman and Stiglitz (1980))

- A theory of "an equilibrium degree of disequilibrium". Noise is essential for existence of equilibrium, because prices are not fully informative and therefore it is valuable for some agent to invest in information production.
- Risk-free asset R and risky asset $u = \theta + \epsilon$. $\theta \sim N(\overline{\theta}, \sigma_\theta^2)$ and $\epsilon \sim N(0, \sigma_\epsilon^2)$
- θ is observable at cost C.
- A fraction λ (endogenous) of ex-ante identical traders choose to pay C and get informed.
- Note that unlike in Grossman (1976) there is no dispersed information.All informed have the same signal. But see Diamond and Verrecchia (1981) and Verrecchia (1982).
- x is supply of risky assets. $x \sim N(\overline{x}, \sigma_x^2)$ is unobservable by uninformed.
- x, θ, ϵ independent.
- Agents start with initial wealth $W_{i0} = \bar{M}_i + \bar{X}_i P_0$. They invest in M_i risk-free dollars and X_i shares of the stock, so that $W_{i0} = M_i + X_i P_0$. At time 1 wealth $W_{i1} = W_{i0} R + X_i(u - RP_0)$
- Agents maximize CARA utility. All with same risk-aversion a. With Gaussian final wealth:

$$\max_{X_i} E[W_{i1}|\mathcal{I}_i] - \frac{a}{2} V[W_{i1}|\mathcal{I}_i] = W_{i0} R + X_i(E[\theta|\mathcal{I}_i] - RP_0) - \frac{a}{2}(V[\theta|\mathcal{I}_i] + \sigma_\epsilon^2).$$

So the optimal demand of each agent is simply:

$$X_i = \frac{(E[\theta|\mathcal{I}_i] - RP_0)}{a(V[\theta|\mathcal{I}_i] + \sigma_\epsilon^2)}.$$

where $\mathcal{I}_i$ depends on the agent's information set.

- Informed agent observe θ so $E[W_{i1}|\theta] = W_{i0}R + X_i(\theta - RP_0)$ and $V[W_{i1}|\theta] = X_i^2\sigma_\theta^2$ and

$$X_I = \frac{\theta - RP_0}{a\sigma_\epsilon^2}$$

- Uninformed agents observe only P_0. What can they learn from price? Market clearing gives $\lambda X_I + (1-\lambda)X_U = x$. It follows that

$$P_0 = \frac{1}{R}(\theta - \psi x) + (1-\lambda)\frac{1}{R}\psi X_u \tag{3}$$

where we define $\psi = \frac{a\sigma_\epsilon^2}{\lambda}$. Thus it is natural to conjecture that uninformed investors, given their own demand, can only infer from price $w_\lambda(\theta, x) = \theta - \psi(x - \overline{x})$ if $\lambda > 0$, otherwise only $w_0(\theta, x) = x$. So we solve the equilibrium as follows:

  (a) conjecture that (all agents conjecture that) $P_0 = \alpha_0 + \alpha_w w_\lambda(\theta, x)$.

  (b) given (a) all agents formulate their optimal demands given their information set. Specifically, $\mathcal{I}_I = \{\theta, P_0\} = \{\theta\}$ and $\mathcal{I}_U = \{P_0\} = \{w_\lambda(\theta, x)\}$

  (c) Plug the optimal demands into the market clearing equation and verify (a).

- Note that if $\lambda > 0$ (there is a typo in GS (A11b)):

$$E[\theta|w_\lambda] = \overline{\theta} + \frac{\sigma_\theta^2}{\sigma_\theta^2 + \psi^2\sigma_x^2}w_\lambda := \overline{\theta} + \beta w_\lambda$$
$$V[\theta|w_\lambda] = \sigma_\theta^2 - \left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \psi^2\sigma_x^2}\right)^2(\sigma_\theta^2 + \psi^2\sigma_x^2) = \frac{\sigma_\theta^2\psi^2\sigma_x^2}{\sigma_\theta^2 + \psi^2\sigma_x^2} = \frac{1}{\frac{1}{\tau_\theta} + \frac{1}{\tau_{\psi x}}}$$

And if $\lambda = 0$

$$E[\theta|w_0] = \overline{\theta}$$
$$V[\theta|w_0] = \sigma_\theta^2$$

Whence the optimal demand of uninformed investors is

$$X_U = \frac{E[\theta|w_\lambda] - RP_0}{a(V[\theta|w_\lambda] + \sigma_\epsilon^2)}$$

- Using X_U and X_I in the market clearing equation confirms the conjecture that P_0 is indeed linear in w_λ and thus we have derived an equilibrium. You can derive explicitly (α_0, α_w) by solving this equation for $P_0(w_\lambda)$:

$$w_\lambda - RP_0 + \overline{x} + (1-\lambda)\psi(\frac{\overline{\theta} + \beta w_\lambda - RP_0}{a(\beta\sigma_\theta^2\psi^2\sigma_x^2 + \sigma_\epsilon^2)}) = 0 \tag{4}$$

- In the case where $\lambda = 0$ then the solution simplifies $P_0(x)$ solves: $\frac{\overline{\theta}-RP_0}{a(\sigma_\theta^2+\sigma_\epsilon^2)} = x$, which gives:

$$P_0 = \frac{\overline{\theta} - a(\sigma_\theta^2 + \sigma_\epsilon^2)x}{R}$$

- Note that unlike in Grossman (1976) the price is not fully revealing even in the presence of informed traders. Uninformed traders can only infer w_λ from prices and would like to know whether prices are high because θ is high or because x is low. This would clearly change their demand for the asset.

  Note that $E[w_\lambda|\theta] = \theta$ and $V[w_\lambda|\theta] = \psi^2\sigma_x^2$. So the equilibrium price signal is more informative the smaller is ψ, i.e., that is the higher λ and smaller is $a\sigma_\epsilon$ (the more there are informed agents and the less they trade for hedging as opposed to informational motives).

  Now we turn to the equilibrium λ.

- A U-agent will want to become I-agent if

$$E[V(R(W_0 - C) + (u - RP)X_I)] \geq E[V(RW_0 + (u - RP)X_U)]$$

  where the 'unconditional' expectations are taken over θ, x, ϵ (ex-ante all investors are identical). Given exponential utility ($V(x) = -e^{-ax}$) we have

$$\frac{E[V(W_{I1})]}{E[V(W_{U1})]} = e^{aC}\frac{E[e^{-aX_I(u-RP)}]}{E[e^{-aX_U(u-RP)}]}$$

Now $E[e^{-aX_I(u-RP)}] = E\left[E[e^{-aX_I(u-RP)}\,|\,\theta, w]\right]$ and

$$\begin{aligned} E[e^{-aX_I(u-RP)}\,|\,\theta, w] &= e^{-aX_I(\theta-RP)+\frac{a^2}{2}X_I^2\sigma_\epsilon^2} \\ &= e^{-\frac{1}{2}\frac{(\theta-RP)}{\sigma_\epsilon^2}} \end{aligned}$$

Similarly, $E[e^{-aX_U(u-RP)}] = E\left[E[e^{-aX_U(u-RP)}\,|\,w]\right]$ and

$$\begin{aligned} E[e^{-aX_U(u-RP)}\,|\,w] &= e^{-aX_U(E[u|w]-RP)+\frac{a^2}{2}X_U^2 V[u|w]} \\ &= e^{-\frac{1}{2}\frac{(E[\theta|w]-RP)^2}{\sigma_\epsilon^2+V[\theta|w]}} \end{aligned}$$

Now, note that $\theta = \overline{\theta} + \beta w + \nu$ where $\nu \perp w$. Thus

$$\begin{aligned} E[e^{-\frac{1}{2}\frac{(\theta-RP)^2}{\sigma_\epsilon^2}}\,|\,w] &= E[e^{-\frac{1}{2}\frac{V[\theta|w]}{\sigma_\epsilon^2}\frac{(\theta-RP)^2}{V[\theta|w]}}\,|\,w] \\ &= \frac{1}{\sqrt{1+\frac{V[\theta|w]}{\sigma_\epsilon^2}}} e^{-\frac{1}{2}\frac{(E[\theta|w]-RP)^2}{\sigma_\epsilon^2+V[\theta|w]}} \end{aligned}$$

where we used that if $z \sim N(\mu_z, 1)$ then $E[e^{-tz^2}] = \frac{e^{-\frac{t\mu^2}{1+2t}}}{\sqrt{1+2t}}$ with $t = \frac{V[\theta|w]}{\sigma_\epsilon^2}$.[2]

Since $V[u|w] = \sigma_\epsilon^2 + V[\theta|w]$ and $V[u|\theta] = \sigma_\epsilon^2$, it follows that

$$E[e^{-aX_I(u-RP)}] = \sqrt{\frac{V[u|\theta]}{V[u|w]}}\,E[e^{-aX_U(u-RP)}]$$

Putting everything together we find:

$$\frac{E[V(W_{I1})]}{E[V(W_{U1})]} = e^{aC}\sqrt{\frac{V[u|\theta]}{V[u|w]}}$$

- An equilibrium fraction of informed agents λ is obtained when at the margin investors are indifferent to becoming informed, thus when

$$e^{aC}\frac{V[u|\theta]}{V[u|w]} = 1$$

[2] $E[e^{-tz^2}] = \int \frac{e^{-tz^2-\frac{1}{2}(z-\mu_z)^2}}{\sqrt{2\pi}}dz = \int \frac{e^{-\frac{1}{2}(1+2t)(z-\frac{\mu}{1+2t})^2-\frac{t\mu^2}{1+2t}}}{\sqrt{2\pi}}dz$ from which the result follows.

It is easy to see that $\gamma'(\lambda) > 0$, thus if $\gamma(0) < 1 < \gamma(1)$ there is a unique 'interior' equilibrium. Else, if $\gamma(0) > 1$, then $\lambda = 0$ is an equilibrium and if $\gamma(1) < 1$ then $\lambda = 1$ is an equilibrium (GS theorem 3).

Note that $\gamma' > 0$ implies that as the fraction of informed increases the relative utility of informed decreases relative to the uninformed (because CARA utility is negative!). So the more informed agents there are in equilibrium the smaller (the more negative) the utility of the informed.

- note that

$$\gamma(\lambda) := e^{aC}\frac{V[u|\theta]}{V[u|w]} = e^{aC}\frac{1+m}{1+m+nm}$$

where $m = \psi^2\frac{\sigma_x^2}{\sigma_\theta^2}$ and $n = \frac{\sigma_x^2}{\sigma_\epsilon^2}$ and further

$$\begin{aligned}\rho^2_{\theta,P} &= \frac{1}{1+m}\\ \rho^2_{\theta,u} &= \frac{n}{1+n}\end{aligned}$$

so the equilibrium determinant of λ is entirely driven by m, n, C, the cost of information and informativeness of the price system (m) and the quality of the information of the informed trader (n).

Specifically the equation can be rewritten as

$$1 - \rho^2_{\theta,P} = \frac{e^{2aC} - 1}{n}$$

we see that

  - increase in the quality of information (n) increases the informational efficiency
  - a decrease in the cost of information, increases informational efficiency and increases the equilibrium λ.
  - a decrease in risk-aversion increase informational efficiency.
  - An increase in σ_x will have no impact on the price efficiency. (why?)

- An equilibrium does not exist without noise (i.e., when $\sigma_x = 0$) if $e^{aC} < \sqrt{1+n}$ (Theorem 5)

  Note that for $\lambda > 0$ then when $\sigma_x = 0$ we have $n, m = 0$ and thus equilibrium would require $e^{aC} = 1$ which cannot hold for $aC > 0$. Now $\lambda = 0$ is not an equilibrium since at that point we have the condition:

$$
\begin{aligned}
\frac{E[V(W_{I1})]}{E[V(W_{U1})]} &= e^{aC}\sqrt{\frac{V[u|\theta]}{V[u|w_0]}} \\
&= e^{aC}\sqrt{\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\epsilon^2}} \\
&= e^{aC}\sqrt{\frac{1}{1+n}} > 1 \ \text{if} e^{aC} > \sqrt{1+n}
\end{aligned}
$$

- When information is perfect ($\sigma_\epsilon = 0$) there never exists and equilibrium. If $\sigma_\epsilon = 0$ then for $\lambda > 0$ we have $nm = 0$ and then $\gamma(\lambda) > 1$ if $aC > 0$. Further, for $\lambda = 0$ we have $\gamma(0) = 0 < 1$.

- Conclusion page 404: GS take issue with the efficient market hypothesis :

  > *"Efficient Markets" theorists have claimed that "at any time prices fully reflect all available information" (see Eugene Fama, p. 383). If this were so then informed traders could not earn a return on their information. We showed that when the efficient markets hypothesis is true and information is costly, competitive markets break down.*
  >
  > [...]
  >
  > *Efficient Markets theorists seem to be aware that costless information is a sufficient condition for prices to fully reflect all available information (see Fama, p. 387); they are not aware that it is a necessary condition. But this is a reducto ad absurdum, since price systems and competitive markets are important only when information is costly (see Fredrick Hayek, p. 452). We are attempting to redefine the Efficient Markets notion, not destroy it. We have shown that when information is very inexpensive, or when informed traders get very precise information, then equilibrium exists and the market price will reveal most of the informed traders' information. However, it was argued in Section III that such markets are likely to be thin because traders have almost homogeneous beliefs.*

The model opens up many interesting theoretical issues (equilbrium traffic-jams, "Efficiently inefficient markets" Garleanu-Pedersen (2020))….) and raises important empirical questions.

- Have markets befome more informationally efficient?
- How can we measure informational efficiency empirically?
- What are the implications for trading volume?
- What are welfare implications of better informational efficiency?
- How to organize markets (microstructure) to achieve such efficiency: OTC, CLOB, market fragmentation…
- How come markets are not fully revealing (I suppose they are not)? If informational efficiency is so important (is it?) for the allocation of resources and welfare, then why not introduce several (non-linear) derivatives, which together would lead to full revelation of the underlying state. Is there something inherent in the trading of several securities that generates additional noise (or market incompleteness) that prevents markets from ever becoming fully revealing?

6 The only game in town Bagehot (1971)

Bagehot (1971) contains the essential idea that investors trading randomly in an efficient market might actually loose on average (which goes against the conventional wisdom of efficient markets of Fama). His beautiful insight is that there are three types of agents in the market : informed speculators who trade with superior information, uninformed agents who trade for 'liquidity-motivated' trades (or may simply be overconfident about their abilities), and market makers whose sole role it is to provide liquidity and transacting whenever opposite orders fail to arrive. Bagehot argues that since market makers will loose systematically to informed agents, she must recoup her losses from trading with uninformed agents, who might thus loose systematically. The market maker does this by setting the bid-ask spread appropriately. With this idea Bagehot offers the fundamental building block for a theory of market liquidity and bid-ask spreads that is purely information-driven (and not based on inventory concerns as in, e.g, Stoll (1978), Ho and Stoll (1981)). We will see two different formalizations of this idea next.

7 Sequential trading in price (or quote)-driven market: Glosten and Milgrom (1985)

7.1 The general framework

- Formalize Bagehot's idea in a model where informed or uninformed agents arrive randomly, *one at a time*, to buy or sell *one* unit from a specialist.

- All orders submitted are market orders. GM do not consider limit orders or the possibility of trading different order sizes. Further, every time a *new* agent is coming to the market, so they cannot strategically delay the decision to trade (their decision is myopic).

- Specialists set bid and ask prices and revise their prices **after** every transaction. Trading is anonymous so specialist cannot distinguish incoming traders. This is a *screening* model. (But see Lee and Wang (2020) for an interesting extension to non-anonymous agents).

- Traders arrive to the market at (random) dates T_k. Arrival rates are not discussed much in the paper, but at a minimum they are not informative about V and/or the identity of the traders. Poisson arrival or deterministic arrival dates seems like what GM have in mind.

- V will be 'realized' at $T_v > t$. It is a random variable with $V > 0$ and $\sigma_V^2 < \infty$. T_v, V independent.

- Informed agents have private information about V.

- All agents get utility $\rho x V + c$ from owning x shares of stock and consuming c. They are risk-neutral but differ in terms of their time preference parameter ρ.

- $\rho = 1$ for market makers. For other traders it is a random variable independent from V and T_v. The randomness in ρ will drive agents desire to trade (gets around the no-trade theorem). It drives agents' private valuations of the asset. $\rho > 1$ they want to buy from the specialist and $\rho < 1$ they want to sell.

- H_t denotes public information. J_t private information. S_t is information of specialist, which contains H_t (GM assume that specialists may be better informed than the public). Assume that A_t, B_t are public information.

- Given that specialist quotes $\{B_t, A_t\}$, it is optimal to buy if $Z_t > A_t$ and sell if $Z_t < B_t$ where

$$Z_t = \rho_t(1 - U_t)E[V_t|J_t, H_t, A_t, B_t] + \rho_t U_t E[V_t|H_t, A_t, B_t]$$

where $U_t = 1$ if an uninformed arrives at t and zero otherwise. Note that ρ_t is a random process whose distribution depends on the identity of the trader arriving to the market .

- GM define F_t to be all the information (it includes H_t, J_t as well as information generated by who arrives at t, i.e., U_t, and information generated by A_t, B_t), so that

$$Z_t = \rho_t E[V|F_t].$$

- Expected profit of specialist is $E[(V - A)\mathbf{1}_{Z_t>A} + (B - V)\mathbf{1}_{Z_t<B} \,|\, S_t]$ which can be rewritten as:

$$(E[V|Z_t > A, S_t] - A)Prob(Z_t > A|S_t) + (B - E[V|Z_t < B, S_t])Prob(Z_t < B|S_t)$$

- Competition among risk-neutral specialists implies the zero-expected-profit conditions:

$$A = E[V|Z_t > A, S_t]$$
$$B = E[V|Z_t < B, S_t]$$

GM assume that there is Bertrand competition at every trade/quote. So specialists set prices so the zero-expected-profit condition holds for *every* trade. If specialists had some market power then it might be optimal to relax this assumption (see Glosten (1989) and Glosten (1994)).

- GM prop 1: We have $A_t > E_t[V] > B_t$. GM allow for S_t to be different from H_t and define E_t to be the expectation with respect to common-knowledge ($H_t \wedge S_t$ which denotes events that are both in S_t and in H_t; typically in many applications it is assumed that $H_t = S_t$).

The proof relies heavily on the law of iterated expectation (and recall that Z_t depends on

$\rho_t, F_t)$:

$$\begin{aligned}
A_t = E_t[A_t] = E_t[A_t|Z_t > A_t] &= E_t[E[V|Z_t > A, S_t]|Z_t > A_t] = E_t[V|Z_t > A_t] \\
&= E_t[E_t[E_t[V|Z_t > A_t, F_t, \rho_t]|Z_t > A, \rho_t]|Z_t > A] \\
&= E_t[E_t[E_t[V|F_t, \rho_t]|Z_t > A, \rho_t]|Z_t > A] \\
&= E_t[E_t[E_t[V|F_t]|Z_t > A, \rho_t]|Z_t > A] \geq E_t[E_t[E_t[V|F_t]|\rho_t]|Z_t > A] \\
&= E_t[E_t[V]|Z_t > A] \\
&= E_t[V]
\end{aligned}$$

where we used the fact that $E[X|X > a] \geq E[X]$ and that ρ_t is independent of V so that $E_t[V|F_t, \rho_t] = E_t[V|F_t]$ and $E_t[E_t[V|F_t]|\rho_t] = E_t[V]$.

- GM prop 2. The sequence of transaction prices $\{p_k\}$ forms a martingale w.r.t. to the specialist's information S_k and the public information H_k, where we define $S_k = S_{T_k^+}$ and $H_k = H_{T_k^+}$ to be the information available to the specialist and the public right after the k^{th} transaction at $T_k < T_V$. Note that after the k^{th} transaction the additional information is the trading price (ask or bid) and the implication that $Z_k < B_k$ or $Z_k > A_k$. The transaction price is

$$\begin{aligned}
A_k \mathbf{1}_{Z_k > A_k} + B_k \mathbf{1}_{Z_k < B_k} &= E[V|S_{T_k}, Z_k > A_k]\mathbf{1}_{Z_k > A_k} + E[V|S_{T_k}, Z_k < B_k]\mathbf{1}_{Z_k < B_k} \\
&= E[V|S_k]
\end{aligned}$$

Since $p_k = E[V|S_k]$ the martingale condition w.r.t S_k-filtration follows from the law of iterated expectation. The martingale condition wrt to H_k follows from the fact that H_k is contained in S_k and since p_k is H_k measurable. Note that while transaction prices are martingales, it does not follow that A_k, B_k are martingales. Indeed, as we show next the Bid-ask spread converges in the limit and thus the difference between $A_k - B_k$ decreases, which implies that A_k, B_k cannot both be martingales.

- GM prop 4: Under some technical condition the expectations of the traders and specialists converge in that $E[V|S_k] - E[V|F_k]$ converges to zero in probability (note that F_k is the information of the kth trader arriving to the market).

- GM's main proposition are quite technical. There paper is in general most well-known for their first example in section 3, and specifically for the binomial example we discuss next.

7.2 Then binomial Glostem-Milgrom model

- Assume $V = v_h$ with prior probability π and $v_l < v_h$ with $1 - \pi$.
- Informed agents know the realization of V and have a $\rho = 1$.
- GM propose 2 models for the preferences of uninformed agents:
  (i) ρ is ∞ (in which case she will buy for sure) with probability γ^b or 0 (in which case she will sell for sure) with probability $\gamma^s = 1 - \gamma^b$.
  (GM use $\gamma^b = 0.5$).
  (ii) ρ is uniformly distributed on $(0, 2)$. This model is less popular so we will consider only the binomial model.
- Assume the proportion of informed traders in the population is α. Assuming independent draws and arrival dates, this implies that the specialist assigns a probablity α that an arriving trader is informed.
- Before any trader arrives the 'fair value' of the asset (given risk-neutrality and zero discounting) is $E(\pi) = \pi v_h + (1 - \pi)v_l$.
- Suppose that at $T_1 < T_v$ a trader arrives to the market and buys at the ask, then with proba $\alpha\pi$ the specialist looses $v_h - A$ to an informed, and with $(1 - \alpha)\gamma^b$ she is earning $A - E$ from an uniformed trader. So the break-even Ask price set by the specialist will satisfy:

$$\alpha\pi(v_h - A) = (1 - \alpha)\gamma^b(A - E)$$

Similarly the bid price will satisfy:

$$\alpha(1 - \pi)(B - v_l) = (1 - \alpha)\gamma^s(E - B)$$

Let's denote these two break-even prices

$$A(\pi) = \frac{\alpha\pi v_h + (1-\alpha)\gamma^b E(\pi)}{\alpha\pi + (1-\alpha)\gamma^b}$$
$$B(\pi) = \frac{\alpha(1-\pi)v_l + (1-\alpha)\gamma^s E(\pi)}{\alpha(1-\pi) + (1-\alpha)\gamma^s}$$

- The A, B just determined form the bid-ask spread before the first trader arrives. Immediately, after the first trade occurs, the specialist will update her quotes to take into account the information disclosed by the trade. If the trade was a buy, then her posterior probability that $\{V = v_h\}$ will increase by Bayes-rule (and commensurately the probability of $V = v_l$ will decrease). The posterior probabilities after the first transaction at T_1 will be:

$$\text{if buy } \pi_{1,b} = P(V = v_h|buy) = \frac{P(buy|V=v_h)P(V=v_h)}{P(buy)} = \frac{(\alpha+(1-\alpha)\gamma^b)\pi}{\pi(\alpha+(1-\alpha)\gamma^b)+(1-\pi)(1-\alpha)\gamma^b} = \frac{(\alpha+(1-\alpha)\gamma^b)\pi}{\alpha\pi+(1-\alpha)\gamma^b}$$
$$\text{if sell } \pi_{1,s} = P(V = v_h|sell) = \frac{P(sell|V=v_h)P(V=v_h)}{P(sell)} = \frac{(1-\alpha)\gamma^s\pi}{\pi(1-\alpha)\gamma^s+(1-\pi)(\alpha+(1-\alpha)\gamma^s)} = \frac{(1-\alpha)\gamma^s\pi}{\alpha(1-\pi)+(1-\alpha)\gamma^s}$$

  So the bid-ask prices prevailing after the first trade (posted prior to the second trade) will be $A(\pi_{1,\tau}), B(\pi_{1,\tau})$ with $\tau = 1$ or 0 depending on whether the first trade was a buy (1) or a sell (0).

- Note that the bid and ask break-even prices we determined actually correspond to the 'no-regret' risk-neutral expected prices, in the sense that:

$$A(\pi) = E[V|buy] = \pi_{1,b}v_h + (1-\pi_{1,b})v_l$$
$$B(\pi) = E[V|sell] = \pi_{1,s}v_h + (1-\pi_{1,s})v_l$$

  which gives an alternative approach to derive the bid-ask prices.

- The process for bid-ask spreads follows then from the sequence of trade arrivals and decisions. The sole state variable is the posterior probability $\pi_{n,\{\tau_1,\tau_2,\dots\tau_n\}}$ which represents the posterior probability that $V = v_h$ after n rounds of trading and having observed the history of trades $\{\tau_1, \dots, \tau_n\}$ where $\tau_i \in \{0, 1\}$ (it equals 1 if the ith trade was a buy and 0 if it was a sell).

Note that Bayes rule gives:

$$\frac{\pi_{1,\tau}}{1-\pi_{1,\tau}} = K_\tau \frac{\pi}{1-\pi}$$

where

$$K_1 = \frac{\alpha + (1-\alpha)\gamma^b}{(1-\alpha)\gamma^b} > 1$$
$$K_0 = \frac{(1-\alpha)\gamma^s}{\alpha + (1-\alpha)\gamma^s} < 1$$

It is then easy to see that the sole state-variable is the number of buys and sells (the order in which these occur do not matter). Specifically if we denote by $n_b = \sum_{i\leq n} \tau_i$, then the bid-ask spread prevailing after the n^{th} trade will be $A(\pi_{n,n_b}), B(\pi_{n,n_b})$ where π_{n,n_b} is obtained from the following equation:

$$\frac{\pi_{n,n_b}}{1-\pi_{n,n_b}} = K_1^{n_b} K_0^{n-n_b} \frac{\pi}{1-\pi}$$

- It is then possible to plot and anlyze the behavior of bid-ask prices and study their convergence as the number of trades increases, as well as their properties with respect to the model parameters (α, γ^b, π). It can in particular be shown that
  - As time goes to infinity prices become fully revealing in the sense that $\lim_{n\to\infty} B(\pi_n) = \lim_{n\to\infty} A(\pi_n) = V$ in probability.
  - The mid price is in general a biased estimate of the expected true value that is in general $\frac{A(\pi)+B(\pi)}{2} \neq E(\pi_n)$.
  - transaction prices are martingales wiht respect to the public information (number of buy and sell trades).
  - Bid and ask prices however are not martingales (wrt to public information), since their difference is a decreaseing process.
  - The bid-ask spread increases with uncertainty about the fundamental value (σ_V), with the fraction of uninformed α.
- The Glosten-Milgrom model (especially in its binomial form) has seen a wide variety of extensions and application (see e.g., the work by Maureen O'hara and co-authors). It has

also led to development of some statistical measures of stock-level market liquidity (such as the PIN and the V-PIN measures of Easley, O'hara, and Paperman (1996) which have been quite controversial, see e.g., Duarte, Hu, and Young (2019).)

8 Batch trading in price-driven market: Kyle (1985)

Kyle (1985) is another very nice formalization of Bagehot's idea. In his model he considers a single risk-neutral informed trader, called the 'insider' (though not necessarily an illegal insider), who has long-lived information about the underlying terminal asset value of the firm. Unlike in GM the agent is strategic about her trading and will optimally dynamically trade over several trading rounds to maximize her expected trading profits. The agent submits only market orders. In addition to the insider there is a continuum of noise traders who trade for non-speculative reasons. A risk-neutral market maker absorbs the total excess demand coming from both types of traders and sets the market clearing price so as to break-even. Since in the Kyle (1985) model informed agents act first, this is more of a signalling model.[3] The model gives insights into (a) the optimal trading strategy of an insider, (b) her trading profits, (c) equilibrium price dynamics, (b) price liquidity (or market depth) as measured by the ubiquitous *Kyle's lambda*. The latter has led to many econometric applications using high-frequency data to measure stock-level trading liquidity.

Kyle considers both a one period model and a multi-auction model where the agent can trade at n trading rounds before her information becomes public. Finally he considers the continuous limit of his model where the agent can trade continuously. We will start with the one-period model.

8.1 The one-period model

-

- The liquidation value $v \sim N(p_0, \Sigma_0)$.

- Noise traders' demand is $u \sim N(0, \sigma_u^2)$ independent of v.

[3]But see the insightful discussion in Brunermeier p. 95,96, who points out that because of the continuous unbounded support of the Gaussian distribution, off-equilibrium strategies need not be considered, and the order in which players act does not seem crucial in the Kyle model, which quite different from standard signalling models.

- The insider observes the realization of v, but not of u, and submits her market order x.[4]

- The risk-neutral market makers observe only total order flow, the net demand from both insider and noise traders, denoted $y = x + u$. Market makers absorb the net demand y at a price so that they break-even. Assuming Bertrand competition the set price so as to have zero expected profits, i.e., $E[-y(v - p_1)|y] = 0$ which implies $p_1 = E[v|y]$.

- An equilibrium then is a trading strategy x that maximizes the expected profits of the insider and a price p that satisfies the zero-expected-profit condition of the market maker:

  (i) $\max_x E[(v - p_1)x|v]$,

  (ii) $p_1 = E[v|y]$.

- To solve this equilibrium Kyle proceeds in three steps: First, conjecture that (the market maker conjecture) the insider chooses a trading strategy of the form $x(v) = \beta(v - p_0)$. Given this conjecture, the market maker will set prices to be linear in order flow:

$$p_1 = E[v|y = \beta(v - p_0) + u] = p_0 + \lambda y$$

  where

$$\lambda = \frac{cov(v, y)}{Var[y]} = \frac{\beta \Sigma_0}{\beta^2 \Sigma_0 + \sigma_u^2}$$

  Second, conjecture that (the insider conjectures) that price responds linearly to order flow, i.e., $p_1 = p_0 + \lambda y$. Given this conjecture the insider will choose her demand to

$$\max_x E[(v - p_1)x] = \max_x E[(v - p_0 - \lambda(x + u))x]$$

  This implies an optimal trading strategy of the insider of the form (from the FOC):

$$x = \frac{v - p_0}{2\lambda}$$

[4] Kyle assumes the insider does not observe noise traders' order flow when submitting her order (Rochet Vila (1994) study the extension where she can and show that the model implications are very similar). Kyle also assumes that the market maker only sees total order flow. Instead, Holden and Subrahmanayam show that if the market maker can see the two order flows separately, without knowing which is informed versus uninformed, the equilibrium is unchanged (in equilibrium the insider's demand 'mimics' the noise trading distribution).

The second order condition for a maximum is satisfied iff $\lambda > 0$.

Lastly, find λ a fixed point so that both conjectures are indeed consistent, i.e, such that:

$$\frac{1}{2\lambda} = \beta$$

plugging in the expression for λ we find that the equilibrium values satisfy:

$$\beta^2 = \frac{\sigma_u^2}{\Sigma_0} \quad \text{and} \quad \lambda = \frac{1}{2\beta} = \frac{\sigma_0}{2\sigma_u}$$

- We can then derive the ex-ante expected profit of the insider:

$$E[(v - p_0)x - \lambda x^2] = \frac{1}{2}\beta^2\Sigma_0 = \frac{1}{2}\sigma_u^2\Sigma_0$$

  and the amount of information that is incorporated into prices due to the information conveyed by order flow:

$$V[v|p_1] = V[v|y] = \Sigma_0 - \lambda^2\sigma_y^2 = \Sigma_0 - \lambda^2(\beta^2\Sigma_0 + \sigma_u^2) = \frac{1}{2}\Sigma_0$$

  The model shows:

  - Prices respond linearly to order flow.
  - Kyle's lambda measures the sensitivity of prices to order flow: it is a signal to noise ratio. The more noise trading the less the adverse selection faced by market makers, the less prices move in reaction to incoming orders.
  - Half of the prior 'fundamental uncertainty' is incorporated into prices with one single trading round.
  - Insider profits are increasing in the amount of prior uncertainty about fundamentals and the amount of noise (which offers camouflage for her trading).

- We now turn to the dynamic multi-auction model to see (i) how insiders will split their orders when they can trade over multiple rounds, and (ii) how this ability to dynamically trade will affect price dynamics and market liquidity. The continuous time limit will be treated below using stochastic calculus.

8.2 The discrete time dynamic model

- Kyle assumes that the insider can trade n times at dates $0 = t_0 < t_1, \ldots, < t_N = T$.
- each round of trading the insider trades Δx_n so that her cumulative position after t auctions is $\sum_{n=1}^{t} \Delta x_n$ starting from $x_0 = 0$.
- The insider maximizes $\max_{\Delta x_n} E[\sum_{n=1}^{n}(v - p_n)\Delta x_n]$
- Noise traders demand is $\Delta u_n \sim N(0, \sigma_u^2 \Delta t)$ to approximate a discrete Brownian motion.
- At each auction round, the market maker absorbs total order flow $\Delta y_n = \Delta x_n + \Delta u_n$ and sets the auction price to break-even: $p_n = E_{n-1}[v|\Delta y_n]$.
- We will seek a linear equilibrium where $p_n = p_{n-1} + \lambda_n \Delta y_n$ and the optimal demand of the insider is of the form $\Delta x_n = \beta_n \Delta t(v - p_{n-1})$.
- The proof is recursive. We start by defining the expected *future* profit value function of the insider $J_n = \max_{\Delta x_t} E[\sum_{t=n+1}^{N}(v - p_t)\Delta x_t]$ and posit that $J_n = \alpha_n(v - p_n)^2 + \delta_n$ for two constants α_n, δ_n which we seek to determine recursively (and with boundary conditions $\alpha_N = \delta_N = 0$).
- By backward induction we can compute the value function of the insider by solving $J_{n-1} = \max_{\Delta x_n} E[(v - p_n)\Delta x_n + J_n]$ recursively from starting from $n = N$.
- At some round n the insider's demand solves $\max_{\Delta x_n} E[(v - p_{n-1} - \lambda_n \Delta x_n)\Delta x_n + \alpha_n(v - p_{n-1} - \lambda_n(\Delta x_n + \Delta u_n))^2 + \delta_n]$
- The first order condition gives: $\Delta x_n = \beta_n \Delta t(v - p_{n-1})$ where

$$\beta_n \Delta t = \frac{1 - 2\alpha_n \lambda_n}{2\lambda_n(1 - \alpha_n \lambda_n)} \quad (\star)$$

- The second order condition gives $\lambda_n(1 - \alpha_n \lambda_n) > 0$.

- Plugging back the first order condition and solving we find that indeed J_{n-1} is quadratic and obtain a recursive solution for $\delta_{n-1}, \alpha_{n-1}$ given by equation 3.15 and 3.16 in Kyle (1985).

- Given the conjecture about the insider's trading strategy that $\Delta x_n = \beta_n \Delta t(v - p_{n-1})$, the market maker will set prices to satisfy:

$$p_n = E_{n-1}[v|\Delta x_n + \Delta u_n] = p_{n-1} + \lambda_n(\Delta x_n + \Delta u_n)$$

  with

$$\lambda_n = \frac{cov_{n-1}[\Delta x_n + \Delta u_n, v]}{V_{n-1}[\Delta x_n + \Delta u_n]} = \frac{\beta_n \Sigma_{n-1}}{\beta_n^2 \Delta t \Sigma_{n-1} + \sigma_u^2} \quad (\star\star)$$

  where we have defined the variance of the fundamental at the beginning of the n^{th} auction $\Sigma_{n-1} = E_{n-1}[(v - p_{n-1})^2]$ (note also that $E_{n-1}[\Delta x_n + \Delta u_n] = 0$).

- By the Gaussian projection theorem the posterior variance of the asset value after the n^{th} auction becomes:

$$\Sigma_n = \Sigma_{n-1} - \lambda_n^2 \Delta t(\beta_n^2 \Delta t \Sigma_{n-1} + \sigma_u^2) = \Sigma_{n-1} - \Sigma_{n-1}\lambda_n \beta_n \Delta t$$

- It is left to show that there exists a unique fixed point (λ_n, β_n) that satisfies the two equations $(\star), (\star\star)$ as well as the second-order condition. Then a numerical solution requires starting at T with boundary conditions $\alpha_N = \delta_N = 0$ and $\Sigma_N = \epsilon$ and iterating back using the difference equations for $\alpha_n, \delta_n, \Sigma_n$ as well as the solutions for β_n, λ_n. An equilibrium is found if the recursion matches the initial condition Σ_0. Kyle shows that one can always find a terminal condition ϵ such that this works.

- We will study the economic implications of the model when we consider the continuous time solution of the Kyle model as derived by Back (1992), which is obtained in closed-form and easy to interpret. In particular, we shall show that in the continuous time limit:
  - Traders will split their orders so as to trade, in expectation, a constant number of shares per unit time.
  - Kyle's lamba is twice as large when the insider can trade continuously.

- All of the private information is ultimately revealed to the market and prices become fully revealing, that is $\Sigma_T = 0$.

- Extensions and follow-up of the Kyle model are numerous and include:
  - Uniqueness of the Kyle equilibrium: existence of non-linear equilibria?
  - Non-Gaussian distribution for the noise and asset value.
  - Multiple competing insiders.
  - Risk-averse insider and residual risk.
  - Activism (the insider can affect the asset value at a cost).
  - Stochstic noise trader volatility.
  - Stochastic horizon.
  - Strategic noise traders.

- Kyle's lambda has become a central empirical measure of stock illiquidity, which can be estimated at various frequencies by regressing price changes (or returns) on order flow. Even the very popular Amihud measure if illiquidity is based on Kyle's insight. See Goyenko, Holden, Trzcinka (2009) for a recent empirical comparison of different liquidity measures.

9 Continuous-time Kyle Model

Here we show the continuous time derivation of the Kyle (1985) model. We follow the derivation in Collin-Dufresne and Fos (2017), who extend the kyle model to allow for stochastic noise trader volatility but retain the assumption that the terminal asset value is normally distributed. This allows to use standard Kalman-Filtering techniques. Instead, Back (1992) allows for arbitrary distribution of the terminal price, which requires more complex techniques. We will see his approach below.

9.1 Continuous time Kyle (1985) model with a finite horizon T

Let's suppose that at T the liquidation value of the firm v will be announced. v is drawn from a prior distrbution $v \sim N(v_0, \Sigma_0)$ at time 0. Only the insider gets to obseve v.

The insider accumulates a total number of shares X_t by choosing an admissible trading rate $\theta_t \in \mathcal{A}$ with $dX_t = \theta_t dt$, so as to maximize the expected value of her trading profits. As shown in Back (1992) it is actually optimal for the insider to trade in an absolutely continuous time fashion prior to the announcement. Indeed, note that, assuming zero interest rates, the dynamics of wealth of the insider is $dW_t = X_t dp_t$. Further $X_T v = X_T P_T = X_0 P_0 + \int_0^T (X_t dp_t + p_t dX_t + d[X,p]_t)$. Thus $W_T = W_0 + \int_0^T X_t dp_t = W_0 - X_0 P_0 + \int_0^T (v - p_t) dX_t - \int_0^T d[X,p]_t$. Since $d[X,p]_t > 0$ in equilibrium, it is optimal to choose an absolutely continuous trading strategy. it follows that the insider will maximize:

$$J(t,p,v) = \max_{\theta_s \in \mathcal{A}} E[\int_t^T (v - p_s)\theta_s ds \mid \mathcal{F}_t^y, v] \tag{5}$$

We define the set of admissible trading strategies $\mathcal{A} = \{\theta_t \ s.t. \ \mathbb{E}[(v - p_t)^2] \leq \infty \ \forall t]\}$.[5] The equilibrium price p_t is set by the competitive risk-neutral market maker so as to break-even on average. Specifically, the zero-profit condition for the market maker implies that

$$p_t = \mathbb{E}[v|\mathcal{F}_t^y] \tag{6}$$

where we denote by $\mathcal{F}_t^y$ the filtration of the market maker generated by observing the cumulative order flow y_t, which is the sum of the informed order flow and noise trading:

$$dy_t = \theta_t dt + \sigma_t dZ_t \tag{7}$$

The cumulative order flow of noise traders is driven by a Brownian motion Z_t with determinstic volatility σ_t.

Thus an equilibrium is defined by a price process p_t and an admissible trading strategy θ_t, that maximizes the profits of the insider in equation (5), while satisfying the market-maker break-even condition equation (6).

To solve the equilibrium, we first conjecture that the trading strategy of the insider is of the form:

$$\theta_t = \beta_t (v - p_t) \tag{8}$$

[5]This technical condition is sufficient to insure that the wealth process of the insider is well-behaved and, in particular, to rule out 'doubling-strategies' as discussed in Dybvig and Huang (1988).

for some deterministic trading speed β_t. Given this conjecture the market maker's filtering problem is a standard conditionally Gaussian problem:

$$
\begin{aligned}
dp_t &= \lambda_t dY_t & (9)\\
\lambda_t &= \frac{\beta_t \Sigma_t}{\sigma_t^2} & (10)\\
d\Sigma_t &= -\lambda_t^2 \sigma_t^2 dt & (11)
\end{aligned}
$$

where $\Sigma_t = \mathbb{E}[(v - p_t)^2 \mid \mathcal{F}_t^y]$ is the conditional posterior variance of the Market maker conditional on observing the continuous order flow.

We can get some intuition for these updating dynamics from the discretized version of the model where we use $dy_t \approx y_{t+dt} - y_t = \beta_t(v - p_t)dt + \sigma_t \epsilon \sqrt{dt}$ with $\epsilon \sim N(0,1)$. Then $p_{t+dt} = E_t[v|dy_t] = E_t[v] + \lambda_t(dy_t - E[dy_t])$ where

$$\lambda_t = \frac{Cov(v, dy_t)}{V(dy_t)} = \frac{Cov(v, dy_t)}{V(dy_t)} = \frac{\beta_t \Sigma_t dt}{\beta_t^2 \Sigma_t dt^2 + \sigma_t^2 dt}$$

Simplifying by dt and taking the limit as $dt \to 0$ gives the expression for λ_t.[6] Further, the posterior variance $\Sigma_{t+dt} = E_t[(v - p_{t+dt})^2 | dy_t] = \Sigma_t - \lambda_t^2 V(dy_t)$. In the limit as $dt \to 0$ we have $V(dy_t) = \sigma_t^2 dt$, which gives the desired equation.

Note that given our conjecture on θ_t, price impact λ_t is itself deterministic. Given the price dynamics in (9) we turn to solving the insider's optimization problem. First, note that his value function can be rewritten as:

$$J(t,p,v) = \max_{\theta_s \in \mathcal{A}} E\left[\int_t^T (v - p_s)\theta_s ds \mid \mathcal{F}_t^y, v\right] \tag{12}$$

The HJB equation is:

$$\max_\theta \left\{ J_t + \frac{1}{2} J_{pp} \lambda_t^2 \sigma_t^2 + J_p \lambda_t \theta + (v - p_t)\theta \right\} = 0. \tag{13}$$

[6]A rigorous derivation of the continuous time Kalman filter equations can be found in Liptser and Shiryaev classic textbook "statistics of Random Processes."

It follows that the first order condition is:

$$J_p \lambda_t + (v - p_t) = 0. \tag{14}$$

We thus guess a quadratic form:

$$J(t,p,v) = \frac{(v-p)^2}{2\lambda_t} + f(t) \tag{15}$$

Using this guess in the HJB equation we find

$$f' - (v-p)^2 \frac{\lambda_t'}{2\lambda_t^2} + \frac{1}{2}\lambda_t \sigma_t^2 = 0 \tag{16}$$

Thus the guess is consistent if:

$$\begin{aligned} f' &= -\frac{1}{2}\lambda_t \sigma_t^2 && (17) \\ \lambda_t' &= 0 && (18) \end{aligned}$$

Thus λ must be constant in equilibrium (otherwise the risk-neutral insider would shift all her demand into states where there is lower price impact). Further, we can solve the equation for $f(t)$ (subject to $f(T) = 0$) to get :

$$\begin{aligned} f(t) &= \frac{1}{2}\lambda \int_t^T \sigma_s^2 ds && (19) \\ &= \frac{\Sigma_t - \Sigma_T}{2\lambda} && (20) \end{aligned}$$

where the second line follows from the explicit solution for the posterior variance given by:

$$\Sigma_t = \Sigma_0 - \lambda^2 \int_0^t \sigma_s^2 ds \tag{21}$$

It remains to pin down the optimal trading strategy of the insider and, in particular, to show there exists a strategy of the form conjectured by the market maker that is indeed optimal for the insider. First, we conjecture that an optimal strategy for the insider must be such that $\lim_{t\to T} p_t = v$, so that there is *no money left on the table.* Otherwise, there would be an incentive for the insider to deviate from the trading strategy (e.g., trade a little more aggressively) to benefit from that price gap.

Thus suppose the insider chooses a strategy as conjectured by the market maker with

$$\beta_t = \frac{\lambda \sigma_t^2}{\Sigma_t}$$

for some constant λ chosen such that $\Sigma_T = 0$, i.e, such that

$$\lambda^2 = \frac{\Sigma_0}{\int_0^T \sigma_s^2 ds} \tag{22}$$

Then from (21) the posterior variance converges to zero at maturity ($\Sigma_T = 0$) which implies that $p_T = v$ (convergence in L^2, which can be shown to imply a.s. convergence). As we show next any strategy with that property is actually optimal for the insider.

Suppose that price is linear in order flow, that is $dp_t = \lambda(\theta_t dt + \sigma_t dZ_t)$. Then consider an arbitrary admissible trading strategy $\theta_t \in \mathcal{A}$ and apply Itô's lemma to the candidate quadratic value function (39):

$$\begin{aligned} J(T, p_T, v) - J(0, p_0, v) &= \int_0^T dJ(t, p_t, v) \\ &= \int_0^T -(v - p_t)(\theta_t dt + \sigma_t dZ_t) \end{aligned}$$

Taking expectation we find that for any admissible trading strategies:[7]

$$J(0, p_0, v) = \mathbb{E}\left[J(T, p_T, v) + \int_0^T (v - p_t)\theta_t dt\right] \tag{23}$$

Now, note that by definition $J(T, p_T, v) = \frac{(v-p_T)^2}{2\lambda} \geq 0$, thus

$$J(0, p_0, v) \geq \mathbb{E}\left[\int_0^T (v - p_t)\theta_t dt\right] \tag{24}$$

[7] The fact that the strategy is admissible guarantees that the stochastic integral is a martingale, since $\mathbb{E}[\int_0^T (v - p_t)^2 \sigma_t^2 dt] < \infty$ for any $\theta_t \in \mathcal{A}$.

for all θ_t.

Further, if we can find an admissible trading strategy such that $p_T = v$ a.s. then $J(T, p_T, v) = 0$ *a.s.* and we obtain an equality in equation (24) which proves the optimality of the strategy. Thus we have shown:

THEOREM 1: *For λ given by equation (22) and Σ given by (21) we have $\lim_{t\to T}\Sigma_t = 0$. Thus an equilibrium exists where the equilibrium price process follows:*

$$dP_t = \kappa_t(v - P_t)dt + \lambda\sigma_t dZ_t \tag{25}$$

$$\kappa_t = \frac{\lambda^2\sigma_t^2}{\Sigma_t} \tag{26}$$

In that equilibrium the informed investor trades as in equation (8) with $\beta_t = \frac{\lambda\sigma_t^2}{\Sigma_t}$.

We see that in the filtration of the insider the price follows a mean-reverting process with mean-reversion rate equal to κ_t. For example we have that $E[p_t - v] = e^{-\int_0^t \kappa_s ds}(p_0 - v)$. Note further that $\frac{d\Sigma_t}{\Sigma_t} = \kappa_t dt$, so that

$$\Sigma_t = \Sigma_0 e^{-\int_0^t \kappa_s ds}.$$

It follows that the expected trading rate of the informed investor in her own filtration is:

$$\mathbb{E}[\theta_t \,|\, v, F_t^Y] = (v - p_0)\beta_t e^{-\int_0^t \kappa_s ds} = \frac{(v - p_0)}{\sqrt{\Sigma_0}} \frac{\sigma_t^2}{\sqrt{\int_0^T \sigma_s^2 ds}} \tag{27}$$

To analyze the implications of our findings it is helpful to define the fundamental uncertainty in terms of its annualized volatility $\Sigma_0 = \sigma_v^2 T$. We see that in equilibrium:

- Price impact, Kyle's lambda, is constant and can always be interpreted as an average signal to noise ratio. When noise trading volatility is constant then $\lambda = \frac{\sigma_v}{\sigma}$. The reason is that since the informed investor is risk-neutral she would otherwise concentrate her trading in those states where she knows for sure that price impact will be lowest. Interestingly this result does not hold when noise trading volatility becomes stochastic (Collin-Dufresne and Fos (2017)).

- The equilibrium price is a martingale in the filtration of the market maker (since $E[dy_t|\mathcal{F}_t^y] = 0$), but it is mean-reverting in the filtration of the insider with mean-reversion rate κ_t with

$\lim_{t \to T} \kappa_t = \infty$. This is a Brownian-bridge type process which converges almost surely to v at maturity T. When noise trading volatility is constant then $\kappa_t = \frac{1}{T-t}$.

- All the information is revealed at T where prices become fully revealing, as $\Sigma_T = 0$. When noise trading volatility is constant $\Sigma_t = \sigma_v^2 (T-t)$.
- The optimal trading strategy of the insider is to trade more aggressively in states where noise trading variance is relatively highest (where she can hide best). When noise trading volatility is constant note that $\beta_t = \frac{1}{\lambda(T-t)}$ which goes to infinity as maturity approaches, so as to profit from any remaining 'inefficiency' ($p_t \neq v$). In expectation the insider trades at a constant rate given by: $\mathbb{E}[\theta_t \,|\, v, F_t^Y] = \frac{(v-p_0)}{\lambda T}$, which has a nice interpretation.
- Price volatility is $\sigma_P(t) = \lambda \sigma_t$ that is more information is revealed to the market in states where noise trading volatility is highest (which is where the insider trades more). Price volatility is constant and equal to $\sigma_P(t) = \sigma_v$ when noise trading volatility is constant (and information is revealed at a constant rate).
- The value function of the insider is $J(0) = \frac{(v-p_0)^2}{2\lambda} + \frac{\Sigma_0}{2\lambda}$ which implies that this unconditional expected profit is given by: $\frac{\Sigma_0}{\lambda} = \sqrt{\Sigma_0 \int_0^T \sigma_t^2 dt}$ which is increasing in both prior fundamental uncertainty and noise trading variance.

9.2 Continuous time Kyle (1985) model with a random poisson distributed horizon T

Let's suppose there is a random announcement time $\tau > 0$ which has a deterministic intensity $\rho_t > 0$. At τ the liquidation value of the firm v will be announced. v is known only to the insider, but has a prior distribution perceived by the market maker of $v \sim N(v_0, \Sigma_0)$.

The insider accumulates a total number of shares X_t by choosing an admissible trading rate $\theta_t \in \mathcal{A}$ with $dX_t = \theta_t dt$ on $t < \tau$, so as to maximize the expected value of her trading profits:

$$J(t,p,v) = \max_{\theta_s \in \mathcal{A}} E[\int_t^\tau (v - p_s)\theta_s \mathbf{1}_{\tau > s} ds \mid \mathcal{F}_t^y, v] \tag{28}$$

We define the set of admissible trading strategies $\mathcal{A} = \{\theta_t \ \ s.t.\ \ \mathbb{E}[(e^{-\int_0^T \rho_s ds} p_T)^2] \leq \infty \ \ \forall T]\}$.[8] The

[8] This technical condition is sufficient to insure that the wealth process of the insider is well-behaved and, in

equilibrium price P_t is set by the competitive risk-neutral market maker so as to break-even on average. Specifically, the zero-profit condition for the market maker implies that

$$P_t = p_t \mathbf{1}_{\tau > t} + v \mathbf{1}_{\tau \leq t}, \tag{29}$$

that is the price jumps on the announcement date to the value v from the *pre-announcement* price p_t given by:

$$p_t = \mathbb{E}[v | \mathcal{F}_t^y, \tau > t] \tag{30}$$

where we denote by $\mathcal{F}_t^y$ the filtration of the market maker generated on $\tau > t$ by observing the cumulative order flow y_t, which is the sum of the informed order flow and noise trading:

$$dy_t = \theta_t dt + \sigma_t dZ_t \tag{31}$$

The cumulative order flow of noise traders is driven by a Brownian motion Z_t with determinstic intensity $\sigma_t < \infty$.

Thus an equilibrium is defined by a pre-anouncement price process p_t and an admissible trading strategy θ_t, that maximizes the profits of the insider in equation (28), while satisfying the market-maker break-even condition equation (30).

To solve the pre-announcement equilibrium, we first conjecture that the trading strategy of the insider (on $\{\tau > t\}$) is of the form:

$$\theta_t = \beta_t (v - p_t) \tag{32}$$

for some deterministic trading speed β_t. Given this conjecture the market maker's filtering problem is a standard conditionally Gaussian problem on the set $\{\tau > t\}$:

$$\begin{aligned} dp_t &= \lambda_t dY_t && (33) \\ \lambda_t &= \frac{\beta_t \Sigma_t}{\sigma_t^2} && (34) \\ d\Sigma_t &= -\lambda_t^2 \sigma_t^2 dt && (35) \end{aligned}$$

where $\Sigma_t = \mathbb{E}[(v - p_t)^2 \,|\, \mathcal{F}_t^y]$ is the conditional posterior variance of the Market maker conditional

particular, to rule out 'doubling-strategies' as discussed in Dybvig and Huang (1988).

on observing the continuous order flow. Note the crucial fact that the announcement date is unpredictable and independent of v, hence knowing $\tau > t$ does not improve the learning of the market maker, i.e., $p_t = \mathbb{E}[v \mid \mathcal{F}_t^y, \tau > t] = \mathbb{E}[v \mid \mathcal{F}_t^y]$.

Note that given our conjecture on θ_t, price impact λ_t is itself deterministic. Given the price dynamics in (33) we turn to solving the insider's optimization problem. First, note that his value function can be rewritten as (on the set $\tau > t$):

$$J(t,p,v) = \max_{\theta_s \in \mathcal{A}} E\left[\int_t^{\infty} e^{-\int_t^s \rho_u du}(v - p_s)\theta_s ds \mid \mathcal{F}_t^y, v\right] \tag{36}$$

The HJB equation is:

$$\max_{\theta}\left\{J_t + \frac{1}{2}J_{pp}\lambda_t^2\sigma_t^2 + J_p\lambda_t\theta - \rho_t J + (v - p_t)\theta\right\} = 0. \tag{37}$$

It follows that the first order condition is:

$$J_p\lambda_t + (v - p_t) = 0. \tag{38}$$

We thus guess a quadratic form:

$$J(t,p,v) = \frac{(v-p)^2}{2\lambda_t} + f(t) \tag{39}$$

Using this guess in the HJB equation we find

$$f' - (v-p)^2\frac{\lambda_t'}{2\lambda_t^2} + \frac{1}{2}\lambda_t\sigma_t^2 - \rho_t\left(\frac{(v-p)^2}{2\lambda_t} + f(t)\right) = 0 \tag{40}$$

Thus the guess is consistent if:

$$0 = f' + \frac{1}{2}\lambda_t\sigma_t^2 - \rho_t f(t) \tag{41}$$

$$\frac{\lambda_t'}{\lambda_t} = -\rho_t \tag{42}$$

Solving the equation for λ we obtain:

$$\lambda_t = \lambda_0 e^{-\int_0^t \rho_u du} \tag{43}$$

Solving the equation for $f(t)$ (subject to $f(\infty) = 0$) gives the solution:

$$f(t) = \lambda_t \int_t^{\infty} e^{-2\int_t^s \rho_u du} \frac{1}{2} \sigma_s^2 ds \tag{44}$$

$$= \frac{\Sigma_t - \Sigma_\infty}{2\lambda_t} \tag{45}$$

Solving for the posterior variance we find:

$$\Sigma_t = \Sigma_0 - \int_0^t \lambda_s^2 \sigma_s^2 ds \tag{46}$$

We can then show the following Theorem:

Theorem 2: *If we can find a constant λ_0 such that $\lim_{t\to\infty} \Sigma_t = 0$ where λ_t, Σ_t are given in equations (43) and (46), then there exists an equilibrium where the price process follows:*

$$dP_t = \kappa_t(v - P_t)dt + \lambda_t \sigma_t dZ_t + (v - P_t) d\mathbf{1}_{\tau \le t} \tag{47}$$

$$\kappa_t = \frac{\lambda_t^2 \sigma_t^2}{\Sigma_t} \tag{48}$$

In that equilibrium the informed investor trades as in equation (32) with $\beta_t = \frac{\lambda_t \sigma_t^2}{\Sigma_t}$. The expected trading rate of the informed investor in her own filtration is:

$$\mathbb{E}[\theta_t \,|\, \tau > t, v, F_t^Y] = (v - p_0)\beta_t e^{-\int_0^t \kappa_s ds} = \frac{(v - p_0)}{\sqrt{\Sigma_0}} \frac{\lambda_0 e^{-\int_0^t \rho_s ds}}{\sqrt{\Sigma_0}} \sigma_t^2 \tag{49}$$

Proof. First we note that if the insider follows the strategy listed in the theorem, then the price $P_t = p_t \mathbf{1}_{\tau > t} + v \mathbf{1}_{\tau \le t}$, where p_t is defined in equation (6). That is the price is consistent with the equilibrium zero-profit condition of the market maker. It remains thus to show that θ_t given in the theorem, is an optimal trading strategy for the insider, i.e., that it solves the optimization problem (12) on $\tau > t$.

To that effect, consider an arbitrary admissible trading strategy $\theta_t \in \mathcal{A}$ and apply Itô's lemma to the candidate quadratic value function (39):

$$
\begin{aligned}
e^{-\int_0^T \rho_s ds} J(T, p_t, v) - J(0, p_0, v) &= \int_0^T e^{-\int_0^t \rho_s ds} \left(dJ(t, p_t, v) - \rho_t J(t, p_t, v) dt \right) \\
&= -\int_0^T e^{-\int_0^t \rho_s ds} (v - p_t)(\theta_t dt + \sigma_t dZ_t)
\end{aligned}
$$

Taking expectation we find that for any admissible trading strategies:[9]

$$
J(0, p_0, v) = \mathbb{E}\left[e^{-\int_0^T \rho_s ds} J(T, p_t, v) + \int_0^T e^{-\int_0^t \rho_s ds} (v - p_t)\theta_t dt \right] \tag{50}
$$

Now, note that by definition $J(T, p_t, v) \geq 0$, thus

$$
J(0, p_0, v) \geq \mathbb{E}\left[\int_0^T e^{-\int_0^t \rho_s ds} (v - p_t)\theta_t dt \right] \tag{51}
$$

for all θ_t and all T. In particular, taking the limit as $T \to \infty$ we have by bounded convergence:

$$
J(0, p_0, v) \geq \mathbb{E}\left[\int_0^\infty e^{-\int_0^t \rho_s ds} (v - p_t)\theta_t dt \right] \tag{52}
$$

Further, if we can find an admissible trading strategy such that $\lim_{T\to\infty} \mathbb{E}\left[e^{-\int_0^T \rho_s ds} J(T, p_T, v) \right] = 0$ then we obtain an equality in equation (52) which proves the optimality of the strategy. Now, note that

$$
\begin{aligned}
\mathbb{E}\left[e^{-\int_0^T \rho_s ds} J(T, p_T, v) \right] &= \mathbb{E}\left[e^{-\int_0^T \rho_s ds} \left\{ \frac{(v - p_T)^2}{2\lambda_T} + f(T) \right\} \right] \\
&= \frac{\Sigma_T}{2\lambda_0} + e^{-\int_0^T \rho_s ds} f(T) \\
&= \frac{2\Sigma_T - \Sigma_\infty}{2\lambda_0}
\end{aligned}
$$

Clearly a sufficient condition for a the right-hand side to go to zero and a strategy to be optimal is

[9]The fact that the strategy is admissible guarantees that the stochastic integral is a martingale, since $\mathbb{E}[\int_0^T e^{-\int_0^t 2\rho_s ds} (v - p_t)^2 \sigma_t^2 dt] < \infty$ for any $\theta_t \in \mathcal{A}$.

that $\lim_{T\to\infty} \Sigma_T = 0$ as stated in the theorem. □

Below we give the explicit solution to the equilibrium when intensity and and noise trading volatility are constant. See Collin-Dufresne, Fos, Muravyev (2019) for an application with increasing intensity and noise-trading volatility to explain the information linkages between equity and option markets. Presentation slides are posted on moodle.

9.2.1 Constant intensity and noise trading volatility

Here we explicitly compute the equilibrium when σ, ρ are both constant.

Solving for the posterior variance and imposing the terminal condition $\lim_{t\to\infty} \Sigma(t) = 0$ we obtain:

$$\Sigma(t) = \frac{\lambda_0^2 \sigma^2}{2\rho} e^{-2\rho t} \tag{53}$$

Then an equilibrium exists if we can find λ_0 such that we satisfy the initial condition $\Sigma(0) = \Sigma_0$. Indeed, we find that the solution is:

$$\lambda_0 = \frac{\sqrt{2\rho\Sigma_0}}{\sigma} \tag{54}$$

wich corresponds to the price impact that would obtain in a finite horizon economy with a fixed horizon of $T = \frac{1}{2\rho}$. The corresponding posterior variance is:

$$\Sigma(t) = \Sigma_0 e^{-2\rho t} \tag{55}$$

Further, we can compute the equilibrium trading strategy:

$$\theta_t = \frac{2\rho e^{\rho t}}{\lambda_0}(v - p_t) \tag{56}$$

and the price process starts from $P_0 = v_0$ and has jump-diffusion dynamics:

$$dP_t = 2\rho(v - P_t)dt + \sqrt{2\rho\Sigma_0} e^{-\rho t} dZ_t + (v - P_t) d\mathbf{1}_{\tau \leq t} \tag{57}$$

We note that the equilibrium price prior to the announcement is a Gaussian mean-reverting process *in the filtration of the insider* with mean-reversion strength equal to twice the announcement intensity and an exponentially decreasing volatility.

We can compute its expectation and variance, conditional on the insider's information:

$$\mathbb{E}_t[p_T - v|v, \tau > T] = e^{-2\rho(T-t)}(p_t - v) \tag{58}$$

$$V_t[p_T - v|v, \tau > T] = e^{-2\rho T}(1 - e^{-2\rho(T-t)})\Sigma_0 \tag{59}$$

And we see that p_t converges in L^2 to v when t goes to infinity.

The expected trading rate becomes

$$\mathbb{E}[\theta_t \,|\, \tau > t, v, F_t^Y] = \frac{(v - p_0)}{\lambda_0 \frac{1}{2\rho}} e^{-\rho t} \tag{60}$$

which starts at time 0 at the same rate one would expect in an economy with a fixed time horizon equal to $1/(2\rho)$ and then decreases at the constant intensity rate.

9.3 Continuous time Kyle (1985) model with stochastic noise trading volatility and stochastic horizon

See presentation slides posted on moodle along with the papers Collin-Dufresne and Fos (2017).

The setup is identical to the one seen in the previous section. The only addition is that now the noise trading volatility process is stochastic driven by it own independent Brownian motion shock W_t. That is:

- The liquidation value of the firm v will be announced at $T.v \sim N(v_0, \Sigma_0)$ at time 0. Only the insider gets to observe v at $t = 0$.

- The insider chooses his trading strategy $dX_t = \theta_t dt$ to maximize $J(t) = \max_{\theta_s \in \mathcal{A}} E[\int_t^T (v - p_s)\theta_s ds \mid \mathcal{F}_t^y, v]$

- Market Makers set price competititely:$p_t = \mathbb{E}[v|\mathcal{F}_t^y]$ where we denote by $\mathcal{F}_t^y$ the filtration of the market maker generated by observing the cumulative order flow y_t (and noise-trading volatility σ_t), which is the sum of the informed order flow and noise trading

$$dy_t = \theta_t dt + \sigma_t dZ_t$$
$$d\sigma_t = m(\sigma, t)dt + \nu(\sigma, t)dW_t$$

- To solve the equilibrium, we first conjecture that the trading strategy of the insider is of the form:

$$\theta_t = \beta_t(v - p_t) \tag{61}$$

for some stochastic trading speed process β_t, which may in particular depend on the noise trading volatility process. Given this conjecture the market maker's filtering problem is a standard conditionally Gaussian problem:

$$\begin{aligned} dp_t &= \lambda_t dY_t && (62) \\ \lambda_t &= \frac{\beta_t \Sigma_t}{\sigma_t^2} && (63) \\ d\Sigma_t &= -\lambda_t^2 \sigma_t^2 dt && (64) \end{aligned}$$

where $\Sigma_t = \mathbb{E}[(v - p_t)^2 \,|\, \mathcal{F}_t^y]$ is the conditional posterior variance of the Market maker conditional on observing the continuous order flow.

- Note that given our conjecture on θ_t, price impact λ_t is itself a stochastic process (possibly depending on t and the history of σ_t).

- Then s Given the price dynamics in (62) we turn to solving the insider's optimization problem. Given the deterministic volatility solution we guess that the value function of the insider is of the form

$$J(t) = J(t, p, v) = \frac{(v - p)^2 + \Sigma_t}{2\lambda_t}$$

Given this guess we not that

$$dJ = -(v - p)(\theta_t dt + \sigma_t dZ_t) + ((v - p)^2 + \Sigma_t)d\frac{1}{\lambda_t} - (v - p)d[p, frac1\lambda_t]_t$$

Integrating and rearranging and taking expectation, we see that if:

(i) Market depth is a martingale, that is $E[d\frac{1}{\lambda_t}] = 0$

(ii) Market depth is independent of the price process $d\frac{1}{\lambda_t}dp_t = 0$.

Then

$$J(0) = E[\int_0^T (v - p_t)\theta_t dt + J(T)]$$

Now, since $J(T) \geq 0$, we see that (i) for any admissible θ we have $J(0) \geq E[\int_0^T (v - p_t)\theta_t dt]$ and (ii) If the exists a θ_t^* such that $p_T = v$ *a.s.* (i.e. there is *no money left on the table*) then $J(0) = E[\int_0^T (v - p_t)\theta_t dt]$, which establishes the optimality of the trading strategy for that price proces.

- It remains to find λ_t, Σ_t so that given a trading strategy pinned down by (63) that is $\beta_t = \frac{\lambda_t \sigma_t^2}{\Sigma_t}$ the price process satisfy the *Bridge property* that $\lim_{t\to T} P_t = v$ *a.s.*. Equivalently we look for a process Σ_t such that $\lim_{t\to T} \Sigma_t = 0$. So we need to solve the forward backward system (on the filtration generated by σ_t, note that this will automatically satisfy condition (ii) above):

  – $E[d\frac{1}{\lambda_t}|\mathcal{F}_t^\sigma] = 0$

  – $d\Sigma_t = -\lambda_t^2\sigma_t^2 dt$ with $\lim_{t\to T} \Sigma_t \to 0$.

- Motivated by the deterministic case, we conjecture that $\lambda_t = \sqrt{\frac{\Sigma_t}{G_t}}$ which decouples the forward backward system:

  – $E[d\sqrt{G_t}|\mathcal{F}_t^\sigma] = -\frac{\sigma_t^2}{\sqrt{G_t}}dt$

  – $\frac{d\Sigma_t}{\Sigma_t} = -\frac{\sigma_t^2}{G_t}dt$ with $\lim_{t\to T} \Sigma_t \to 0$.

  The solution is then

  $$\Sigma_t = \Sigma_0 e^{-\int_0^t \frac{\sigma_t^2}{G_t}dt}$$

  where G_t solves the backward (recursive) equation

  $$\sqrt{G_t} = E[\int_t^T \frac{\sigma_t^2}{\sqrt{G_t}}dt]$$

- Collin-Dufresne and Fos (2017) establish under what conditions a solution to G exists such that $\Sigma_T = 0$ and $p_T = v$ *a.s.*. This implies tha the equilibrium exists with $\lambda_t = \sqrt{\frac{\Sigma_t}{G_t}}$ where the price process follows:

$$dP_t = \kappa_t(v - P_t)dt + \lambda_t \sigma_t dZ_t \tag{65}$$

$$\kappa_t = \frac{\sigma_t^2}{G_t} \tag{66}$$

In that equilibrium the informed investor trades as in equation (61) with $\beta_t = \frac{\kappa_t}{\lambda_t}$.

We see that in this model

  - Price impact is in general stochastic that it tends to increase on average (because its inverse, market depth) is a martingale. Liquidity deteriorates over the trading day in equilibrium so that the insider is willing to trade early and give up her option to delay trading to wait for better liquidity states. (Stochastic noise trading volatility gives the insider a liquidity timing option.)
  - The optimal trading strategy of the insider is to trade more in states where price impact it lowest as well as in states where noise trading volume is highest (relative to the average level of uninformed noise trading).
  - Price volatility is stochastic driven by noise trading volatility. In states where noise trading volatility is high, the insider trades more and more information gets into prices.

9.4 Continuous time Kyle (1985) with non-normally distributed terminal value: Back (1992)

- Given a price function $P(t, Y_t)$, insider maximizes

$$\max_{\theta} E\left[\int_0^T (v - P(t, Y_t))\theta_t \, dt \mid v\right]. \tag{67}$$

- Market Maker has prior $v \sim F(x) = Prob(v \leq x)$ (not necessarily normal!) and observes total order flow Y_t:

$$dY_t = \theta_t dt + \sigma dZ_t$$

where Z_t is standard Brownian motion and where $dX_t = \theta_t dt$ is the trading strategy of the insider. Note that it is not a priori obvious why the insider should trade in an absolutely continuous fashion (i.e., without any dZ_t term). However, given that the model assumes that σ is common-knowledge any additional Brownian component would effectively be 'observed' by the market maker and lead to price impact that would reduce the profits of the insider. See the discussion page 393-394 in Back (1992). Here we will proceed with the conjecture that the insider's strategy is of that form.

- An equilibrium is a pair (P, θ) s.t. trading strategy θ maximizes (67) given P and

$$P(t, Y_t) = E\left[v \mid \mathcal{F}_t^Y\right] \tag{68}$$

- Search for an equilibrium where $P(y,t)$ only depends on aggregate order flow. Then, risk-neutrality of the market maker implies $P(Y_t, t) = \mathbb{E}[h(Y_T) \,|\, \mathcal{F}_t^Y]$ for some function $h(Y)$.

- Note that in continuous time it is natural to assume that the insider can effectively 'observe' total order flow Y_t and hence also σZ_t and thus she can condition her trades on the noise trader demand. This is because in equilibrium Y_t can be inverted from $P(Y_t, t)$. This is different from the discrete Kyle model where it is assumed that the insider cannot condition her trades on noise trader demand which it does not observe. But see Rochet and Vila (1992) who study a one-period Kyle model, where the insider can observe noise trader demand.

- If the insider's trading strategy is unpredictable (i.e. $E[\theta_t|\mathcal{F}_t^y] = 0$), then

$$P(y,t) = \mathbb{E}[h(y + \sigma(Z_T - Z_t)) \,|\, \mathcal{F}_t^Z].$$

  So the price function is pinned down by $h(\cdot)$ (given law of Z).

- If the insider leaves no-money on the table at maturity, then $h(Y_T) = v\;\; a.s.$.

- To find candidate $h(\cdot)$, we use the fact that $Y_T \sim N(0, \sigma^2 T)$. So we need $h^{-1}(v) \sim N(0, \sigma^2 T)$, which implies $Prob(h^{-1}(v) \leq y) = N(\frac{y}{\sigma\sqrt{T}})$ where $N(x)$ is the standard normal cumulative distribution function. But we also have by definition $Prob(v \leq h(y)) = F(h(y))$. Thus our candidate

$$h(y) = F^{-1}(N(\frac{y}{\sigma\sqrt{T}}))$$

. Note that $h'(y) > 0$ (since $N'(x) > 0$ and $F'(x) > 0$).

- It remains to show that (a) there exists a trading strategy θ_t such that y_t is a Brownian motion on its own filtration that converges almost surely to $h^{-1}(v)$, and (b) that this strategy is optimal for the insider.

- The insider maximizes

$$E\left[\int_0^T (v - P(t, Y_t))\theta_t \, dt \mid v\right]$$

- HJB equation for value function $J(Y_t, t)$ is linear in control θ. Obtain:

$$\begin{array}{lrcl} (HJB) & 0 & = & \frac{1}{2} J_{YY}\sigma^2 + J_t \\ (FOC) & 0 & = & J_Y + v - P(Y, t) \end{array}$$

- Using Feynman-Kac we seek (a 'no-trade') solution of the form $J(y, t) = \mathbb{E}_t[g(y + \sigma(Z_T - Z_t))|v]$ for some function $g(\cdot)$.

- To determine g we assume by 'continuity' that the second (FOC) equation holds at T (where $P(Y, t) = h(Y)$). So g should satisfy $g'(y) + v - h(y) = 0$, which leads us to guess a function of the form $g(y) = sup_{\bar{y}} \int_y^{\bar{y}} (v - h(z))dz$. Optimizing over $\bar{y}$ leads to the condition $\bar{y} = h^{-1}(v)$ and our guess

$$g(y) = \int_y^{h^{-1}(v)} (v - h(z))dz$$

Clearly we have that $g(y)$ satisfies:

$$0 = g_Y + v - h(y)$$

- It follows (with sufficient regularity so we can take derivatives inside the expectation) that for our candidate value function:

$$J_Y(y,t) = \mathbb{E}_t[g_Y(y+\sigma(Z_T - Z_t))|v] = \mathbb{E}_t[h(y+\sigma(Z_T - Z_t)) - v|v] = P(y,t) - v$$

Thus, the candidate J satisfies the HJB equation and FOC.

- It follows that for any θ_t:

$$J(Y_0,t) = E_0\left[g(Y_T) + \int_0^T (v - P(Y_t,t))\theta_t dt\right]$$

- Now note that $g(y) \geq 0$ and that for the particular strategy θ_t^* such that $Y_T^* = h^{-1}(v)$ a.s. (which is equivalent to $h(Y_T^*) = v$) we have $g(Y_T^*) = 0$. This establishes the optimality of our candidate value function, if indeed we can find the strategy with the desired convergence property.

- To complete the proof it remains to show that we can find a strategy θ_t^* such that $dY_t = \theta_t^* dt + \sigma dZ_t$ converges a.s. to $h^{-1}(v)$ at T and such that $E[\theta_t^*|\mathcal{F}_t^Y] = 0$. Note that the process $dY_t = \frac{1}{T-t}(h^{-1}(v) - Y_t)dt + \sigma dZ_t$ is a Brownian bridge that converges almost surely to $h^{-1}(v)$ at maturity in the filtration of the insider. Further we have $dY_t = \sigma dZ_t^y$ where Z_t^y is a standard Brownian motion on F_t^Y. Thus the insider's trading strategy

$$dX_t = \frac{1}{T-t}(h^{-1}(v) - Y_t)dt$$

is optimal.

- To prove the 'Brownian bridge property,' note that $\epsilon = h^{-1}(v) \sim N(0,\sigma^2 T)$. Consider $dY_t = \beta_t(\epsilon - Y_t)dt + \sigma dZ_t$ for some constant c to be the observation equation for the random variable ϵ and define $p_t = E[\epsilon|\mathcal{F}_t^Y]$ and $\Sigma_t = E[(\epsilon - p_t)^2|\mathcal{F}_t^Y]$. Then standard Gaussian

Kalman filtering gives

$$
\begin{aligned}
dp_t &= \frac{\beta_t \Sigma_t}{\sigma^2}(dY_t - \beta_t(p_t - Y_t)dt) \\
&= \frac{\beta_t^2 \Sigma_t}{\sigma^2}(\epsilon - p_t)dt + \frac{\beta_t \Sigma_t}{\sigma^2}\sigma dZ_t \\
d\Sigma_t &= -\frac{\beta_t^2 \Sigma_t^2}{\sigma^2}dt
\end{aligned}
$$

Set $\beta_t = \frac{1}{T-t}$, to obtain the following solution :

$$
\begin{aligned}
dp_t &= \frac{1}{T-t}(\epsilon - p_t)dt + \sigma dZ_t \\
\Sigma_t &= \sigma^2(T-t)
\end{aligned}
$$

Given initial condition $p_0 = Y_0 = 0$ we see that with $\theta_t = \frac{1}{T-t}(\epsilon - Y_t)$ we have $p_t = Y_t \ \ \forall t \ a.s.$. This establishes that Y_t is a martingale on its own filtration ($E[\theta_t] = 0$) that converges almost surely to ϵ at T ($\Sigma_T = 0$).

9.4.1 The Gaussian case

Suppose that v is normally distributed with mean v_0 and variance $\sigma_v^2 T$. Then $F(x) = N(\frac{(x-v_0)}{\sigma_v\sqrt{T}})$, thus $F^{-1}(y) = v_0 + \sigma_v\sqrt{T}N^{-1}(y)$. It follows that:

$$h(y) = F^{-1}(N(\frac{y}{\sigma\sqrt{T}})) = v_0 + \lambda y$$

where

$$\lambda = \frac{\sigma_v}{\sigma}$$

This implies $P(y,t) = E[h(y + \sigma(Z_T - Z_t)] = v_0 + \lambda y$ which implies $dP_t = \lambda dY_t$. Further $h^{-1}(v) = \frac{v-v_0}{\lambda}$ so $\theta_t = \frac{1}{\lambda(T-t)}(v - v_0 - \lambda Y_t)dt = \frac{1}{\lambda(T-t)}(v - P_t)dt$ and we obtain the same equilibrium as in the continuous time Kyle model

9.4.2 The Log-Normal case

Suppose that $\log v$ is normally distributed with mean v_0 and variance $\sigma_v^2 T$. Then $F(x) = Prob(v \leq x) = Prob(\log v \leq \log x) = N(\frac{(\log x - v_0)}{\sigma_v \sqrt{T}})$, thus $F^{-1}(y) = \exp(v_0 + \sigma_v \sqrt{T} N^{-1}(y))$. It follows that:

$$h(y) = F^{-1}(N(\frac{y}{\sigma\sqrt{T}})) = \exp(v_0 + \lambda y)$$

where

$$\lambda = \frac{\sigma_v}{\sigma}$$

This implies $P(y,t) = E[h(y + \sigma(Z_T - Z_t)] = \exp(v_0 + \lambda y + \frac{1}{2}\lambda^2\sigma^2(T-t))$ which implies $\frac{dP_t}{P_t} = \lambda(dY_t - \frac{1}{2}\lambda\sigma^2 dt)$. Further $h^{-1}(v) = \frac{\log v - v_0}{\lambda}$ so $\theta_t = \frac{1}{\lambda(T-t)}(\log v - v_0 - \lambda Y_t)dt = \frac{1}{\lambda(T-t)} \log \frac{v}{P_t} dt$ and we obtain 'log-normal' version of the continuous time Kyle model, where prices are log-normally distributed (consistent with limited liability and empirical return features) and it makes sense to think of an illiquidity 'Kyle-lambda' computed from returns as opposed to price changes.

10 Kyle and Lee When are financial markets strategic?

- Kyle (1989) and Kyle and Lee (2018) provide nice equilibrium models where agents act strategically, in the sense that they take into account the impact of their trading on the equilibrium price. In so doing, they avoid the problem of the rational expectation equilibrium (REE) discussed above in the context of Grossman and Grossman-Stiglitz's models. The issue of trader 'skizophrenia', who learn from prices the information of other traders even though they assume that they themselves have no impact on price. The fact that in equilibrium (if noise is sufficiently small) there may be no incentive to acquire information, since prices reveal that information 'costlessly' and therefore free-riding is optimal. If agents act strategically, then their trading will be less aggressive thus changing the information acquisition incentives. The papers discuss whether and under what conditions the strategic equilibrium approaches the competitive REE. The conclusion is that while prices may look similar in the two economies in many conditions (e.g., when the number of traders increases), quantities rarely look alike (trading behaviors remain quite different). Further, the incentives with respect to information acquisition or the organisation of markets are very different across the two equilibrium concepts. Lastly, the papers show that measures based on Kyle-lambda type of price impact are only imperfect measures of market illiquidity as they do not properly take into account the impact of adverse selection on the actual quantity traded in equilibrium. On a technical side, Kyle (1989) uses a model with exogenous noise traders, whose demand for shares is perfectly price inelastic, whereas Kyle and Lee (2018) uses trader specific hedging demands driven by endowment shocks to model noise. The latter is more suitable to think about welfare implications (every trader has a well-defined utility) and strategic trading equilibria (every trader is strategic). In a nutshell, Kyle (1989) is the strategic version of the Grossman-Stiglitz (1980) REE model, whereas Kyle-Lee (2018) is the strategic version of the Diamond and Verrecchia (1981) REE model.

- Asset value $v \sim N(0, \sigma_v^2)$.
- $n = 1, \dots N$ agents with CARA utility and absolute risk-aversion A.
- Each agent observes a signal $i_n = v + e_n$ where $e_n \sim N(0, \frac{\sigma_v^2}{\tau_I})$. All signals are independent. Note that τ_I measures the precision of the signal relative to the prior uncertainty.[10]
- Each agent receives a certain number of shares of the asset as a random endowment $s_n \sim N(0, \sigma_s^2)$.
- Each agent submits *symmetric linear demand schedules* $X(P, i_n, s_n) = -\pi_p P + \pi_I i_n - \pi_s s_n$.[11]
- Market clearing $\sum_n X(P, i_n, s_n) = 0$.
- Study 2 kinds of equilibrium:
  (i) Strategic (Bayesian-Nash) equilibrium: each agent chooses her demand curve to maximize her expected utility, taking as given all other $(N-1)$ agents demand curve.
  (ii) Competitive REE: each agent chooses her demand curve to maximize her expected utility taking the price functional as given.

We focus first on the strategic equilibrium

[10] Kyle (1989) also has M uninformed agents (without signal).

[11] The paper focuses only on symmetric linear equilibria and does not investigate whether their exists non-linear equilibria.

10.1 Strategic equilibrium

- If agent n is *strategic* then she considers the *residual supply curve*

$$P = \underbrace{\frac{\pi_I}{\pi_p} i_{-n} - \frac{\pi_s}{\pi_p} s_{-n}}_{p_n} - \underbrace{\frac{1}{(N-1)\pi_p}}_{\lambda} X_n$$

where we define $s_{-k} = \frac{1}{N-1} \sum_{n \neq k} s_n$.

- So the agent can infer $\frac{\pi_p}{\pi_I} p_n = v + e_{-n} - \frac{\pi_s}{\pi_I} s_{-n} := v + e^P$ from the market price, which is a noisy estimate of the other traders information. note that

$Var[e^p] = \frac{\sigma_v^2}{(N-1)\tau_I} + \frac{\pi_s^2}{\pi_I^2 (N-1)} \sigma_s^2.$

- Agent n maximizes $J = \mathbb{E}[(X_n + s_n)v - X_n P \mid P, i_n, s_n] - \frac{A}{2} \mathrm{Var}[(X_n + s_n)v \mid P, i_n, s_n]$

- FOC gives

$$X_n = \frac{(\mathbb{E}[v \mid P, i_n, s_n] - A s_n \mathrm{Var}[v \mid P, i_n, s_n] - p_n)}{2\lambda + A \mathrm{Var}[v \mid P, i_n, s_n]}$$

- The SOC is $2\lambda + A\mathrm{Var}[v \mid P, i_n, s_n] > 0$.

- Note that

$$\begin{aligned}
\mathbb{E}[v \mid P, i_n, s_n] &= \mathbb{E}[v \mid \frac{\pi_p}{\pi_I} p_n, i_n, s_n] \\
&= \frac{1}{1 + \tau_I + (N-1)\tau_I \varphi} ((N-1)\tau_I \varphi \frac{\pi_p}{\pi_I} p_n + \tau_I i_n)
\end{aligned}$$

where we define

$$\varphi = \frac{\sigma_v^2}{(N-1)\tau_I V[e^P]}$$

and

$$\hat{V} := \mathrm{Var}[v \mid P, i_n, s_n] = \frac{\sigma_v^2}{1 + \tau_I + (N-1)\tau_I \varphi}$$

- The proof of this relies on the multi-variate Gaussian projection theorem. It follows also straightforwardly from the result that if one observes n signals $s_i = v + \epsilon_i$ with independent

$\epsilon_i \sim N(0, \sigma_i^2)$, then the minimum variance portfolio of the signals $S = \sum_i w_i s_i = v + \sum_i w_i \epsilon_i$ with weights $w_i = \frac{\tau_i}{\sum_j \tau_j}$ and precisions $\tau_i = \frac{1}{\sigma_i^2}$, is a sufficient statistic for the n signals in the sense that $v = \beta S + \nu$ with $\beta = \frac{Cov(v,S)}{V(S)} = \frac{\sigma_v^2}{\sigma_v^2 + \sum_i w_i^2 \sigma_i^2}$ satisfies $\nu \perp s_i \;\; \forall i$.

- Plug the solutions into the FOC to obtain :

$$(2\lambda + A\hat{V})X_n = \frac{1}{1 + \tau_I + (N-1)\tau_I \varphi}((N-1)\tau_I \varphi \frac{\pi_p}{\pi_I} p_n + \tau_I i_n) - A s_n \hat{V} - p_n$$

- Now we use the definition $p_n = P + \lambda X_n$ to obtain

$$(\lambda + A\hat{V})X_n = \frac{1}{1 + \tau_I + (N-1)\tau_I \varphi}((N-1)\tau_I \varphi \frac{\pi_p}{\pi_I}(P + \lambda X_n) + \tau_I i_n) - A s_n \hat{V} - P$$

rearranging

$$((1+\tau_I)\lambda + A\sigma_v^2)X_n = \tau_I i_n - A\sigma_v^2 s_n - (1 + \tau_I + (N-1)\tau_I \varphi(1 - \frac{\pi_p}{\pi_I}))P$$

- Matching with the assumed form $X_n = -\pi_p P + \pi_I i_n - \pi_s s_n$ we get three equations for π_p, π_s, π_I:

$$\begin{aligned}
((1+\tau_I)\lambda + A\sigma_v^2)\pi_p &= 1 + \tau_I + (N-1)\tau_I \varphi(1 - \frac{\pi_p}{\pi_I}) \\
((1+\tau_I)\lambda + A\sigma_v^2)\pi_I &= \tau_I \\
((1+\tau_I)\lambda + A\sigma_v^2)\pi_s &= A\sigma_v^2 \\
\lambda &= \frac{1}{(N-1)\pi_p}
\end{aligned}$$

from which it follows that

$$\begin{aligned}
\frac{\pi_p}{\pi_I} &= (1 + \tau_I^{-1} + (N-1)\varphi(1 - \frac{\pi_p}{\pi_I})) \\
\frac{\pi_s}{\pi_I} &= \frac{A\sigma_v^2}{\tau_I}
\end{aligned}$$

and finally:

$$\frac{\pi_p}{\pi_I} = 1 + \frac{1}{\tau_I + (N-1)\tau_I\varphi}$$
$$\frac{\pi_s}{\pi_I} = \frac{A\sigma_v^2}{\tau_I}$$
$$((1+\tau_I)\lambda + A\sigma_v^2)\pi_p = \tau_I(1 + \frac{1}{\tau_I + (N-1)\tau_I\varphi})$$

- Solving for π_p, π_s, π_I we get the expression for the optimal demand X_n of each agent. Imposing market clearing $\sum_{n=1}^N X_n = 0$ we obtain the market clearing price

$$P = \frac{\pi_I}{\pi_p}\sum_n i_n - \frac{\pi_s}{\pi_p}\sum_n s_n$$

- substituting the expression for λ into the SOC we get a necessary condition for an equilibrium to exist.

- Putting all the results together we get proposition 1.

10.2 The competitive equilibrium

- Assume that all agents have demand functions of the form $X_n = \pi_I i_n - \pi_s s_n - \pi_P P$.

- When agents take prices as given market clearing implies $\sum_n X_n = 0$, that is $P = \frac{\pi_I}{\pi_p}\bar{i} - \frac{\pi_s}{\pi_p}\bar{s}$ where we define $\bar{s} = \frac{1}{n}\sum_n s_n$.

- Of course, given her own signal and endowment (i_n, s_n) the agent can infer from the price the same signal as in the competitive case $\frac{\pi_p}{\pi_I}P - (i_n - \frac{\pi_s}{\pi_I}s_n) = v + e_{-n} - \frac{\pi_s}{\pi_I}s_{-n} := v + e^P$ from the market price, which is a noisy estimate of the other traders information. note that $Var[e^p] = \frac{\sigma_v^2}{(N-1)\tau_I} + \frac{\pi_s^2}{\pi_I^2(N-1)}\sigma_s^2$.

- Agent n maximizes $J = \mathbb{E}[(X_n + s_n)v - X_n P \mid P, i_n, s_n] - \frac{A}{2}\mathrm{Var}[(X_n + s_n)v \mid P, i_n, s_n]$

- FOC gives
$$X_n = \frac{(\mathbb{E}[v \mid P, i_n, s_n] - As_n\mathrm{Var}[v \mid P, i_n, s_n] - P)}{A\mathrm{Var}[v \mid P, i_n, s_n]}$$

- The SOC is $A\mathrm{Var}[v \mid P, i_n, s_n] > 0$ which is always satisfied in this case!

- Note that

$$\mathbb{E}[v \mid P, i_n, s_n] = \frac{1}{1+\tau_I+(N-1)\tau_I\varphi}((N-1)\tau_I\varphi\frac{\pi_p}{\pi_I}P + \tau_I i_n)$$

where we define

$$\varphi = \frac{\sigma_v^2}{(N-1)\tau_I V[e^P]}$$

and

$$\hat{V} := \mathrm{Var}[v \mid P, i_n, s_n] = \frac{\sigma_v^2}{1+\tau_I+(N-1)\tau_I\varphi}$$

- Plug the solutions into the FOC to obtain :

$$(A\hat{V})X_n = \frac{1}{1+\tau_I+(N-1)\tau_I\varphi}((N-1)\tau_I\varphi\frac{\pi_p}{\pi_I}P + \tau_I i_n) - As_n\hat{V} - P$$

rearranging we get:

$$A\sigma_v^2 X_n = \tau_I i_n - A\sigma_v^2 s_n - (1+\tau_I+(N-1)\tau_I\varphi(1-\frac{\pi_p}{\pi_I}))P$$

- Matching with the assumed form $X_n = -\pi_p P + \pi_I i_n - \pi_s s_n$ we get three equations for π_p, π_s, π_I:

$$\begin{aligned}
(A\sigma_v^2)\pi_p &= 1+\tau_I+(N-1)\tau_I\varphi(1-\frac{\pi_p}{\pi_I}) \\
(A\sigma_v^2)\pi_I &= \tau_I \\
(A\sigma_v^2)\pi_s &= A\sigma_v^2
\end{aligned}$$

which are the same equations as in the strategic case with $\lambda = 0$! It follows that

$$\pi_s = 1$$
$$\pi_I = \frac{\tau_I}{A\sigma_v^2}$$
$$\frac{\pi_p}{\pi_I} = 1 + \frac{1}{\tau_I + (N-1)\tau_I\varphi}$$

- Solving for π_p, π_s, π_I we get the expression for the optimal demand X_n of each agent. Imposing market clearing $\sum_{n=1}^{N} X_n = 0$ we obtain the market clearing price

$$P = \frac{\pi_I}{\pi_p}\sum_n i_n - \frac{\pi_s}{\pi_p}\sum_n s_n$$

  Since the ratios $\frac{\pi_I}{\pi_p}$ and $\frac{\pi_s}{\pi_p}$ are unchanged the equilibrium price is actually identical relative to the strategic case!

- Putting all the results together we get proposition 2. Note the interesting results:
  - Equilibrium always exists in the competitive case, but not in the strategic case (requires $N > 2$ and SOC).
  - The equilibrium price function is the same in both models. This implies that in equilibrium, in the competitie model there is a residual demand curve with a strictly positive slope (λ), even though the informed agent assumes it is zero! (the schizophrenia).
  - The quantities traded by the agents differ in that

$$X_n^{Strat} = \chi X_n^{Comp}$$

    where

$$\chi = (\frac{N-2}{N-1} - 2\varphi)/(1-\varphi)$$
$$\varphi = \frac{\tau_I}{\tau_I + A^2\sigma_v^2\sigma_s^2}$$

    Agents trade less in the strategic equilibrium. Kyle and Lee define χ as the measure of market competitiveness. When $\chi = 1$ both equilibria are identical, strategic considerations become insignificant. Note that χ depends only on N and on the *informational efficiency* φ and decreases in both. For strategic markets to become competitive requires both $N \to \infty$ and $\varphi \to 0$. The latter means that either their signal becomes uninformative or their hedging motive dwarfs their information motive of trading.

> Summarizing, financial markets become perfectly competitive if and only if there are infinitely many traders and relative informational efficiency ' approaches zero, in which case hedging completely dominates speculation. As long as speculation remains important in financial markets, markets remain imperfectly competitive.

11 Illiquidity, stock returns, and the role of market making: Grossman-Miller (1988)

Propose a model of market making and liquidity based on the temporal imbalance of order flow and demand and supply for immediacy. It is different from the information based theory with risk-neutral market makers (Glosten-Milgrom (1985)), and more related to the inventory- models where risk-averse market makers hold risky inventory and provide immediacy to arriving customers. GM point out that market makers' supply of immediacy stands in contrast with thin, illiquid markets (such as housing) where immediacy is not a primary concern and issues regarding moral hazard and adverse selection lead the 'market makers' to offer services in marketing, advertising, search services rather than providing immediacy by acting as principal traders.

- Model has three CARA agents with same risk-aversion a. Agent 1 arrives in period 1 and wants to sell i units of an asset (e.g., because she is endowed with i units of that asset). Agent 2 arrives in period 2 and wants to buy i units of the same asset (e.g., she has a risky endowment of $-i$ units). The asset pays a risky dividend in period 3 which is $P_3 \sim N(\mu, \sigma^2)$. If the two agents arrived to the market at the same time, they would trade with each other and there would be zero trade imbalance. Instead, there are M market makers who are present in the market at all times and will provide immediacy, buying from agent 1 in period

1 and selling to agent 2 in period 2.

- Assuming zero risk-free rate, agent 1 starts with B_0 cash and i shares to invest in x_1 shares at time 1 so that $W_1 = x_1P_1 + B_1 = iP_1 + B_0$. Then at time 2 she rebalances her portfolio so that $W_2 = x_2P_2 + B_2 = x_1P_2 + B_1$. Finally $W_3 = x_2P_3 + B_2$. Combining we find that $W_3 = x_2(P_3 - P_2) + x_1(P_2 - P_1) + iP_1 + B_0$ which can be rewritten as:

$$W_3 = B_0 + \hat{x}_1(P_2 - P_1) + \hat{x}_2(P_3 - P_2) + iP_3$$

  with $\hat{x}_i = x_i - i$ defining the trade net of the initial exposure.

- Proceeding recursively and using the CARA-normal setup $\max_{\hat{x}_2} E_2[U(W_3)] = \max_{\hat{x}_2} E_2[W_3] - \frac{a}{2}V_2[W_3]$ which gives:

$$\hat{x}_2 = \frac{E_2[P_3] - P_2}{aV_2[P_3]} - i := D(i)$$

- Now at $t = 2$ the newly arriving agent 2 has offsetting excess demand $\hat{x}_2^{new} = D(-i)$, and the M market makers each demand $D(0)$, thus market clearing (sum of excess demands is zero) gives:

$$D(i) + MD(0) + D(-i) = 0$$

  Together this implies

$$P_2 = E_2[P_3]$$

  and thus $\hat{x}_2^* = -i = -\hat{x}_2^{new}$.

- It follows that for agent 1, $W_3 = B_0 + \hat{x}_1(P_2 - P_1) + \hat{x}_2^*(P_3 - P_2) + iP_3 = B_0 + \hat{x}_1(P_2 - P_1) + iP_2$. Thus at date 1, agent 1 seeks x_1 to $\max_{\hat{x}_1} E_1[W_3] - \frac{a}{2}V_1[W_3]$ which gives:

$$\hat{x}_1 = \frac{E_1[P_2] - P_1}{aV_1[P_2]} - i = \frac{E_1[P_3] - P_1}{aV_1[E_2[P_3]]} - i$$

  Note that we need some news about the terminal payoff to arrive in period 2 (so that $V_1[P_2] = V_1[E_2[P_3]] > 0$). for the model to display some interesting properties.

- Similarly the demand from market makers in period 1 will be $x_1^m = \frac{E_1[P_2] - P_1}{aV_1[P_2]} = \frac{E_1[P_3] - P_1}{aV_1[E_2[P_3]]}$.

Thus market clearing implies : $\hat{x}_1 + Mx_1^m = 0$ which implies

$$\frac{E_1[P_2] - P_1}{aV_1[P_2]} = \frac{i}{1+M} \equiv x_1^m$$

- Define the excess return earned by market makers:

$$E_1[r] = E_1[\frac{P_2}{P_1}] - 1 = x_1^m P_1 a V_1[r]$$

The excess return earned by market makers is increasing in the 'dollar' inventory $x_1^m P_1$ held by market makers, if and only if $V_1[r] > 0$. The latter requires $V_1[P_2] = V_1[E_2[P_3]] > 0$. That is, there needs to be some risk in the time-2 prices, which in the context of the model means some news about the terminal payoff P_3 must be disclosed at time 2. Else, holding the inventory from period 1 to period 2 is not risky, and there is no role for providing immediacy.

- How much of the desired customer trade is completed at time 1 versus time 2 (this is an interesting alternative measure of liquidity that is different from the classic bid-ask spread)?

$$x_1^* = x_m - i = -\frac{M}{M+1}i \quad \text{and} \quad x_2^* = -i = x_1^* - \frac{i}{M+1}.$$

Thus only a fraction $\frac{M}{M+1}$ of the desired trade $(-i)$ is done in period 1. The larger the number of market makers M, the more liquid the market. Because of limited risk-bearing capacity of market makers, it is not optimal to execute all of the desired trade rightaway, but rather to hold on to some exposure.

- What is the equilibrium number of market makers?

Assume that it costs C to become one. Then free entry implies M solves:

$$E[U(B_0 - C + x_1^m(P_2 - P_1) + x_2^m(P_3 - P_2)] = E[U(B_0)]$$

using the fact that $x_2^m = 0$ and $x_1^m = \frac{i}{M+1}$ we obtain the following condition:

$$E[e^{-a[-C + \frac{i}{M+1}(P_2 - P_1)}] = 1$$

where $P_2 = E_2[P_3]$. Assuming that i, P_2 are independent and normally distributed we obtain:

$$E[e^{aC+\frac{1}{2}(\frac{ai}{M+1})^2 V_1[P_2]-a\frac{i}{M+1}E_1[P_2-P_1]}] = E[e^{aC-\frac{1}{2}(\frac{ai}{M+1})^2 V_1[P_2]}] = 1$$

Assuming that i is normally distributed we can compute the expectation using the moment generating function of a Chi-Squared random variable to obtain:

$$\frac{1}{\sqrt{1+t}} e^{-\frac{E[i]^2}{1+t}\frac{t}{2}} = e^{-ac}$$

where

$$t = \frac{a^2 V_1[P_2]}{(1+M)^2} V[i]$$

If $E[i] = 0$ (imbalances are zero on average) then t is determined only by ac and is in fact increasing in ac. So M is larger the smaller ac and the larger the risk to hedge ($V[i]V[P_2]$).

- Implication of the model for price auto-correlation (an alternative measure of illiquidity, relative to Bid-ask spread also advocated by Roll (1984)). Define autocorrelation to be

$$q = \frac{Cov(P_2 - P_1, P_1 - E_0[P_1])}{\sqrt{V[P_2 - P_1]V[P_1 - E_0[P_1]]}}$$

- Use $P_2 = E_2[P_3]$ and $P_1 = E_1[P_2] - \frac{i}{1+M} a V_1[P_2]$ and $E_0[P_1] = E_0[P_2]$ (i.e., assuming $E_0[i] = 0$) to get that $P_2 - P_1 = P_2 - E_1[P_2] + \frac{i}{1+M} a V_1[P_2]$ and $P_1 - E_0[P_1] = E_1[P_2] - \frac{i}{1+M} a V_1[P_2] - E_0[P_2]$ Further, assume that $s^2 = V_1[P_2] \equiv V_1[P_2 - E_1[P_2]] = V[E_1[P_2] - E_0[P_2]]$

  So that

$$q = -\frac{t}{\sqrt{1+t}}$$

  This implies the autocorrelation in price changes is negative and determined solely by the cost of becoming a market maker. More negative autocorrelation expected the more costlier it is to become a market maker (the lower the risk-bearing capacity of market makers).

- Nice discussion of different market structures p. 620-622, where they compare (i) highly liquid stocks, where there is a lot of order-flow and where the specialist can play the role of an "auctionneer", (ii) for smaller stocks role of the specialist becomes more prevalent, there are fewer market makers who need to take the stocks on their books (where the minimum tick size rule plays a role in preventing exercise of monopoly power) and where illiquidity can arise if too large an order comes to the market, (iii) the upstairs market which functions like a search market for larger "block-size" trades, (iv) the pure OTC search markets, with no obligation to maintain any continuous market making presence. Suggests many interesting questions: What is the 'optimal' market structure? Does it depend on security characteristics? Does it depend on client/customer characteristics and needs? What is interaction between market structure, liquidity and efficiency?

- Discussion of the relation to the literature on Bid-ask spreads and on the role of market makers in the 1987 Krash. In general, the idea of this paper is that Market maker risk-bearing capacity and fluctuations therein can explain (time-varying) market liquidity in ways that simply looking at bid-ask spreads may be mis-leading. Also, this is a different source of illiquidity than the information cost discussed previously. See also the interesting discussion of GM by Whitcomb, posed on moodle.

12 Illiquidity and Stock Returns: Amihud Mendelson (1986)

Why do investors hold different portfolios? How does illiquidity affect asset returns? AM envision that both investors demand is shaped by different liquidity needs (they have different trading horizons, maybe are subject to different liquidity shocks). Further assets have different liquidity characteristics (their bid-ask spreads differ, which might reflect a cost of immediacy or informational frictions). AM propose a theory where in equilibrium different liquidity clienteles choose portfolios with different liquidity characteristics, and where in equilibrium the expected return on individual assets will be increasing in the bid-ask spread. That is, the cross-section of stocks' expected returns will be affected by illiquidity characteristics.

> *Our model predicts that higher-spread assets yield higher expected returns, and that there is a clientele effect whereby investors with longer holding periods select assets with higher spreads. The resulting testable hypothesis is that asset returns are an increasing and concave function of the spread. The model also predicts that expected returns net of trading costs increase with the holding period, and consequently higher-spread assets*

yield higher net returns to their holders. Hence, an investor expecting a long holding period can gain by holding high-spread assets.

AM also test their theory empirically and find support for it:

> *We test the predicted spread-return relation using data for the period 1961-1980, and find that our hypotheses are consistent with the evidence: Average portfolio risk-adjusted returns increase with their bid-ask spread, and the slope of the return-spread relationship decreases with the spread. Finally, we verify that the spread effect persists when firm size is added as an explanatory variable in the regression equations. We emphasize that the spread effect is by no means an anomaly or an indication of market inefficiency; rather, it represents a rational response by an efficient market to the existence of the spread.*

AM's paper is one of the first to link microstructure with more traditional asset pricing and suggest that illiquidity could be 'priced,' and affect the cross-section of expected returns. It is interesting that their mechanism is not related to the traditional risk-premium explanation (unlike the Acharya-Pedersen 'liquidity risk-premium' model we will see next). Further, their model is one of different clienteles, i.e., effectively of segmented markets, where in equilibrium, there is not necessarily one common pricing kernel that prices all assets consistently (arbitrage is ruled out, because shorting is not allowed). In their own words:

> *This study highlights the importance of securities market microstructure in determining asset returns, and provides a link between this area and mainstream research on capital markets. Our results suggest that liquidityincreasing financial policies can reduce the firm's opportunity cost of capital, and provide measures for the value of improvements in the trading and exchange process. In the area of portfolio selection, our findings may guide investors in balancing expected trading costs against expected returns. In sum, we demonstrate the importance of market-microstructure factors as determinants of stock returns.*

12.1 The theory

- We consider a slightly simpler version of the AM model first.

- Investor A is risk-neutral and has random trading horizon τ with intensity λ_A.

- Each security pays continuous dividend δ and is in finite supply (e.g., 1 unit). Note that AM assume that each security pays different dividend stream δ_i.

- Each security quoted by market makers at fixed bid-ask spread $[P_i - C_i\delta, P_i]$. Note that AM use a proportional spread instead i.e., assume that the Bid is $(1 - S_i)P_i$.

- Exogenous risk-free rate r, infinite supply, no trading costs.

- No short-sales.

- Equilibrium price is $P_i(0) = \mathbb{E}[\int_0^\tau e^{-rt}\delta dt + e^{-r\tau}(P_i(\tau) - C_i\delta)]$, so that the expected return on asset i is equal to the risk-free rate and the risk-neutral agent is indifferent between investing in the risky asset or holding the risk-free asset.

- Note that at any time t, (dropping the i subscript for simplicity) we have $P_t\mathbf{1}_{\tau>t} = \mathbb{E}_t[\int_t^\infty e^{-r(s-t)}\delta\mathbf{1}_{\tau>s}ds + \int_t^\infty e^{-r(s-t)}(P_{s^-} - C\delta)d\mathbf{1}_{\tau\leq s}]$ so that $e^{-rt}P_t\mathbf{1}_{\tau>t} + \int_0^t e^{-rs}\delta\mathbf{1}_{\tau>s}ds + \int_0^t e^{-rs}(P_{s^-} - C\delta)d\mathbf{1}_{\tau\leq s}$ is a martingale. Using the fact that $E_t[d\mathbf{1}_{\tau\leq t}] = \lambda\mathbf{1}_{\tau>t}dt$ and $E_t[d\mathbf{1}_{\tau>t}] = -\lambda\mathbf{1}_{\tau>t}dt$ and assuming that 'no-common-jumps' (i.e., $dP_t d\mathbf{1}_{\tau>t} = 0$) we get the expression:

$$\mathbf{1}_{\tau>t}\left\{dP_t - (r+\lambda)P_t dt + \delta dt + (P_t - C\delta)\lambda dt\right\} = 0$$

  which implies :

$$\frac{dP_t + \delta(1 - C\lambda)dt}{P_t} = r dt$$

  We recognize on the LHS the expected return on the stock at time t conditional on having not sold prior to t. Looking for a stationary solution with a constant $P_t = P$, we easily solve for P (note that we could also directly solve the expectation for $P_i(0)$ above under the assumption of a constant P_i).

- Solution $P_i = \frac{\delta}{r}(1 - \lambda_A C_i) =: \overline{P}_i^A$

$\Rightarrow$ If A is marginal holder of security i then it trades at a discount $D_i^A = \lambda_A C_i$ to the friction-less value, that accounts for the NPV of expected future transaction costs.

- Note:
  - $D_1^A = \lambda_A C_1$ increasing in C_1 $\quad\rightarrow$ liquidity premium effect.
  - $\frac{D_1^A}{C_1} = \lambda_A$ is decreasing in $\mathbb{E}[\tau_A] = \frac{1}{\lambda_A}$ $\quad\rightarrow$ liquidity clientele effect.

- Consider now what happens if there are two clienteles $\lambda_A < \lambda_B$ (that is A has a longer average horizon than B) and two types of assets $C_1 < C_2$.

- If A has unlimited capital she will hold both assets (since $\overline{P}_i^A > \overline{P}_i^B$ $\forall i = 1, 2$ because A has a longer horizon)

$\Rightarrow$ There should be no liquidity clientele effect (only a liquidity premium).

$\rightarrow$ Clientele effects should be more prevalent when funding is restricted (crisis?).

- If A has limited capital and cannot buy all the bonds, then we expect that B will be marginal in bond 1 (the low cost asset, which is B's comparative advantage since B has a shorter horizon). So $P_1 = \overline{P}_1^B$.

- Further, since A must choose not to buy (all of) security 1 at this price, A must earn more than the risk-free rate on security 2 in equilibrium. So we expect $P_2 < \overline{P}_2^A$.

- Indeed, in equilibrium A must be indifferent between security 1 and 2:

$$\mathbb{E}[\frac{dP_2 + \delta dt}{P_2} - \frac{C_2 \delta \lambda_A}{P_2} dt] = \mathbb{E}[\frac{dP_1 + \delta dt}{P_1} - \frac{C_1 \delta \lambda_A}{P_1} dt] > r dt$$

- This implies that $P_2 = \frac{1 - C_2 \lambda_A}{1 - C_1 \lambda_A} P_1 = \frac{\delta}{r}(1 - D_2)$ where $D_2 = 1 - \frac{1 - C_1 \lambda_B}{1 - C_1 \lambda_A}(1 - C_2 \lambda_A)$

- It is then easy to show that $\overline{D_2^B} \equiv C_2 \lambda_B > D_2 > C_2 \lambda_A \equiv \overline{D_2^A}$. This implies that B does not want to hold the high t-cost asset 2 in equilibrium, wheras A is indifferent between holding asset 2 or asset 1 and earns a return higher than the risk-free rate for holding the most illiquid assets.

- It is easy to show that in equilibrium:
  - $D_2 > D_1$ (liquidity premium effect). Securities with higher transaction costs will have higher gross expected returns as $\frac{\delta}{P_2} = \frac{r}{1 - D_2} > \frac{r}{1 - D_1} = \frac{\delta}{P_1}$.

- $\frac{D_2}{C_2} < \frac{D_1}{C_1}$ (liquidity clientele effect). Controlling for the security specific trading cost (i.e., the measured bid-ask spread) high t-cost firms have lower expected "gross returns". This is because higher t-cost firms are held by longer horizon investors who effectively amortize the t-cost over a longer horizon. So the selection mechanism implies that gross-returns seem less sensitive to t-costs in the cross-section.
- $\frac{D_2}{\lambda_A C_2} > \frac{D_1}{\lambda_B C_1}$ (*clientele equilibrium rents effect*). Since long-horizon investors need to be compensated to hold higher transaction costs assets (and not flock to the lower transaction cost assets which are held by the short-horizon investors), they earn extra-rents on these assets in equilibrium.

- In addition this model has implications for the cross-section of portfolio holdings:
  - Credit (Funding) Market conditions should affect the empirical results: clientele effects should be stronger when funding market conditions are tight.
  - In equilibrium long-horizon investors are indifferent between high and low liquidity assets (so might expect their portfolios to be less informative than those of short-horizon investors).
  - Long-horizon investors should extract rents in equilibrium. So one should see higher average returns net of trading costs for them.

    This could potentially be tested by normalizing spreads by the expected transaction costs (turnover $\times$ bid-ask spread) of the marginal investor.

- The AM theory relies on strong assumptions (risk-neutrality, no short-sales, limited funding resources, exogenous T-costs and exogenous trading horizon...)
- The AM theory is entirely about exogenous trading costs and exogenous horizon. But what determines asset illiquidity (the C_i parameters)?
- AM actually derive a richer model with M securities and N investors and show that:

  > *The ensuing equilibrium has the following characteristics: (i) market-observed average returns are an increasing function of the spread: (ii) asset returns to their*

holders, net of trading costs, increase with the spread: (iii) there is a clientele effect, whereby stocks with higher spreads are held by investors with longer holding periods: and (iv) due to the clientele effect, returns on higher-spread stocks are less spread-sensitive. giving rise to a concave return-spread relation.

:

In particular, they prove the two propostions:

- Proposition I (clientele effect). Assets with higher spreads are allocated in equilibrium to portfolios with (the same or) longer expected holding periods.
- Proposition 2 (spread-return relationship). In equilibrium, the observed market (gross) return is an increasing and concave piecewise-linear function of the (relative) spread.

They proceed to test this using empirical data.

12.2 Empirical test

- Use CRSP data.
- Measure illiquidity using as proxy for the trading cost the average of the beginning and end-of-year relative bid-ask spreads for each of the years 1960-1979.
- Divide the data into twenty overlapping periods of eleven years each, consisting of a five-year estimation period E_n, a five-year portfolio formation period F_n, and a one-year cross-section test period T_n $(n = 1, 2, ..., 20)$.
- For each E_n estimate CAPM beta.
- For each F_n rank all stocks in seven portfolios based their end of F_n-year bid-ask spread. Then within each of the seven portfolios rank on their E_n-β and split into seven portfolios. This procedure results in $7 \times 7 = 49$ equal-sized portfolios ranked on bid-ask spread and beta.
- Then for each of the 49 portfolios estimate their CAPM beta and their average bid-ask spread during the E_n period.

- Lastly they run a pooled time-series and crossectional regression of the average monthly excess portfolio returns R^e_{pn} in each T_n periods on the portfolio β_{Pn} and their bid-ask spreads S_{Pn} as in equation (8) in their paper.

- Results summarized in section 3.3. lend support to their theory. They find in particular:

  > *The coefficient of S_{Pn} implies that a 1% increase in the spread is associated with a 0.211% increase in the monthly risk-adjusted excess return. The coefficient of β declines when the spread variable is added to the equation, indicating that part of the effect which could be attributed to β may, in fact, be due to the spread.*

13 Illiquidity and Stock Returns: Acharya Pedersen (2005)

AM (1986) showed theoretically and empirically that illiquidity can affect the cross-section of measured stock returns. AP (2005) present a simple model where illiquidity *risk* will be priced and carry a risk-premium. Their focus in on liquidity risk, i.e., the fact that illiquidity may change over time in ways that agents care about and thus may want to hedge against. In equilibrium, liquidity betas will be priced characteristics of stock returns. Their idea is a simple extension of the CAPM logic and indeed in their model the CAPM holds exactly, but for returns defined net of (exogenously specified) t-costs.

13.1 The Model

- Overlapping generations of agents who live for 2 periods. Agent $n = 1, \ldots, N$ of generation born at time t has CARA utility with risk-aversion A_n, is endowed with $e_n(t)$, trades at time t, and then sells all of her shares at time $t+1$ (to generation $t+1$) in order to consume at $t+1$.

- There are a total of I risky securities with total supply of S_i $\forall i = 1, \ldots, I$ of each risky security. Each security pays a dividend $D_i(t+1)$, it has an illiquidity cost of $C_i(t+1)$ (when sold) and has price $P_i(t)$. $C_i(t)$ and $D_i(t)$ follow stochastic processes:

$$C_i(t+1) = \bar{C} + \rho_C(C_i(t) - \bar{C}) + \eta_i(t+1)$$
$$D_i(t+1) = \bar{D} + \rho_D(D_i(t) - \bar{D}) + \epsilon_i(t+1)$$

where $\epsilon_i(t), \eta_i(t)$ are jointly normally distributed with zero mean and variance matrices $V[\epsilon] = \Sigma_D, V[\eta] = \Sigma_C, E[\epsilon\eta^\top] = \Sigma_{DC}$ and covariance matrices and uncorrelated across time.

- Short-selling is not allowed.
- There is a risk-free asset with rate of return $R_f > 1$ in infinite supply (a risk-free technology).
- Assuming that generation t-agent n's wealth will be conditionally normally distributed, the CARA assumption implies that she chooses the vector of risky asset holdings y in order to $\max_y E[W_{t+1}] - \frac{A_n}{2} V[W_{t+1}]$, where terminal wealth is given by

$$W_{t+1} = (e_t - y^\top P_t)R_f + y^\top(D_{t+1} + P_{t+1} - C_{t+1})$$

- The first-order conditions implie:

$$y_n(t) = (A_n V[D_{t+1} + P_{t+1} - C_{t+1}])^{-1} (E_t[D_{t+1} + P_{t+1} - C_{t+1}] - P_t R_f)$$

Market Clearing implies that in equilibrium $\sum_n y_n(t) = S$, which implies a market clearing price vector:

$$P_t = \frac{1}{R_f} (E_t[D_{t+1} + P_{t+1} - C_{t+1}] - AV[D_{t+1} + P_{t+1} - C_{t+1}])$$

where the aggregate risk-aversion is $A^{-1} = \sum_n \frac{1}{A_n}$

We look for a stationary equilibrium of the form $P_t = p_0 + p_d D_t - p_C C_t$. Note that :

$$
\begin{aligned}
E_t[D_{t+1} + P_{t+1} - C_{t+1}] &= p_0 + E_t[D_{t+1}(1 + p_d) - (1 + p_c)C_{t+1}] \\
&= p_0 + (1 + p_d)(\bar{D} + \rho_D(D(t) - \bar{D})) - (1 + p_c)(\bar{C} + \rho_C(C(t) - \bar{C})) \\
V_t[D_{t+1} + P_{t+1} - C_{t+1}] &= V[(1 + p_D)\epsilon_{t+1} - (1 + p_C)\eta_{t+1}]
\end{aligned}
$$

So the coefficients must satisfy

$$
\begin{aligned}
R_f p_d &= (1 + p_d)\rho_D \\
R_f p_C &= (1 + p_c)\rho_D \\
R_f p_0 &= p_0 + (1 + p_d)(1 - \rho_D)\bar{D} - (1 + p_C)(1 - \rho_C)\bar{C} - AV[(1 + p_D)\epsilon_{t+1} - (1 + p_C)\eta_{t+1}]
\end{aligned}
$$

- With this solution we see that indeed W_{t+1} is normally distributed. It follows that the CAPM holds in every period (every agent will optimally choose to hold a long position in risky assets consisting of a fraction $\frac{A}{A_n}$ of the market portfolio S (this follows from comparing the market clearing condition and the FOC) and a long position in the risk-free asset. In particular, note there is no shorting in equilibrium. Thus for any asset i its net return satisfies the CAPM equation:

$$
\begin{aligned}
E_t[R_i(t+1) - c_i(t+1)] &= R_f + \beta_i(E_t[R_M(t+1) - c_M(t+1)] - R_f) \\
\beta_i &= \frac{Cov_t[R_i(t+1) - c_i(t+1), R_M(t+1) - c_M(t+1)]}{V_t[R_M(t+1) - c_M(t+1)]} \\
R_i(t+1) &= \frac{P_i(t+1) + D_i(t+1)}{P_i(t)} \\
c_i(t+1) &= \frac{C_i(t+1)}{P_i(t)} \\
R_M(t+1) &= \sum_{i=1}^{N} \omega_i(t) R_i(t+1) \\
\omega_i(t) &= \frac{S_i P_i(t)}{\sum_j S_j P_j(t)} \\
c_M(t+1) &= \sum_{i=1}^{N} \omega_i(t) c_i(t+1)
\end{aligned}
$$

- AP derive 2 additional propositions that show that (i) the conditional expected excess return on a portfolio is typically increasing in its illiquidity when illiquidity is persistent ($\rho_C > 0$) that is persistence in illiquidity implies return predictability, and (ii) the covariance between a portfolio's return and its cost of trading is typically negative.

- Expanding the CAPM equation we get the main equation (8) of the AP paper, which expresses measured excess returns in terms of the illiquidity premium (similar to AM1986), a traditional CAPM market beta, and **three additional illiquidity betas**. It is this equation that they proceed to test empirically.

- A good summary of their findings is on page 377 and 378. Based on average rate of turnover the main component of the risk-premium due to liquidity seems to be the AM86 liquidity level premium (estimated at 3.5% per year). The three liquidity risk-premium component contribute an additional 1.1% per year to a stock's measured "gross" risk-premium. The most significant beta seems to be the $Cov(c_i, R_M)$ beta, which contributes 0.86% of the liquidty risk-premium and suggests investors care to hold securities who are liquid when the market returns are low.

- They test the 4-beta CAPM in five steps, see page 11

14 Differences in Beliefs and short-sale constraints: Miller (1977)

So far we considered models where agents are differentially informed, but share the same priors so that if they had access to the same signals they would come to the same conclusion (i.e., posterior). Instead we will now consider a series of papers where agents have different priors and 'agree to disagree' (see Aumann (1976) for a theoretical discussion of the impossibility for rational agents with common-knowledge to 'agree to disagree').

Miller (1977) is one of the first to realize that combining differences in beliefs with short-sale constraints can explain many economic phenomena,[12] including:

- why stocks with highest risk have low returns (the volatility anomaly of Ang, Hodrick, Xing, Zhang (1996)),
- the poor long run results on new issues of stocks (the new issuance anomaly of Daniel adn Titman (2006)),
- the presence of discounts from net value for closed end investment companies (the closed-end fund discount of Lee, Schleifer, and Thaler (1991)), and
- the lower than predicted rates of return for stocks with high systematic risk (the betting against beta anomaly of Black (1978) and Frazzini-Pedersen (2014)).

His idea rests on the simple insight that if investors have enough resources to buy all the shares of a company and there are short-sales constraints, then the equilibrium valuation will reflect the beliefs of the most optimistic investors. Thus stocks will tend to be overvalued. This effect will be more prevalent amongst stocks which are riskier (where the differences in beliefs are largest). Thus riskier stocks will tend to be more overvalued and underperform. New issues and younger stocks are more likely to present large disagreement, etc....

To illustrate his basic idea in a simple framework we present a static model from Kerry Back's textbook. Then we will consider a dynamic setting due to Harrison and Kreps (1978) and a continuous time setting due to Scheinkman and Xiong (2011).

[12] Note that this is Edward Miller and not Merton Miller!

15 The static Model

This simple model is based on Back's chapter18.4 description of the Chen, Hong and Stein (2002) paper.

- A continuum of CARA agent of mass 1 indexed by h with risk-aversion coefficient γ trade one risky asset with payoff $P_1 \sim N(\mu_h, \sigma^2)$.

- There exists a risk-free rate R_f.

- Each agent's demand is $x_h = \frac{\mu_h - P_0 R_f}{\gamma\sigma^2}$

- Each agent is endowed with 1 unit of stock.

- Assume μ_h is uniformly distributed on $(\mu^* - \Delta, \mu^* + \Delta)$

- In the absence of short-sales then the aggregate demand of all investors is

$$D_m = \frac{1}{2\Delta} \int_{\mu - \Delta^*}^{\mu^* + \Delta} \frac{\mu - P_0 R_f}{\gamma\sigma^2} d\mu = \frac{\mu^* - P_0 R_f}{\gamma\sigma^2}$$

  and market clearing imposes that $D_m = 1$. It follows that the unconstratined equilibrium price is

$$P_0^{unc} = \frac{\mu^* - \gamma\sigma^2}{R_f}$$

- Suppose instead that there are short-sales constraints. Then investor h's optimal demand becomes : $x_h = \frac{\mu_h - P_0 R_f}{\gamma\sigma^2} \mathbf{1}_{\mu_h \geq P_0 R_f}$

- If the unconstrained price is such that the short-selling constraint will never bind for any investor, then the market clearing price will be unchanged. This will be the case if $\mu_h \geq \mu^* - \gamma\sigma^2 \;\; \forall h$, or equivalently if $\Delta \leq \gamma\sigma^2$ (that is if there is not much dispersion in beliefs).

- Consider now the case where $\Delta > \gamma\sigma^2$ then there is a positive mass of investors (with $\mu_h < \mu^* - \gamma\sigma^2$) who would like to short the stock and cannot. Thus the equilibrium demand is:

$$
\begin{aligned}
D_m &= \frac{1}{2\Delta}\int_{\mu-\Delta^*}^{\mu^*+\Delta} \frac{\mu - P_0R_f}{\gamma\sigma^2}\mathbf{1}_{\mu\geq P_0R_f}d\mu \\
&= \frac{1}{2\Delta}\int_{P_0R_f}^{\mu^*+\Delta} \frac{\mu - P_0R_f}{\gamma\sigma^2}d\mu \\
&= \frac{(\mu^* + \Delta - P_0R_f)^2}{4\Delta\gamma\sigma^2}
\end{aligned}
$$

- Then market clearing imposes $D_m = 1$ which gives the market clearing price:

$$P_0 = \frac{1}{R_f}(\mu^* + \Delta - 2\sqrt{\Delta\gamma\sigma^2})$$

- Note in particular that the price with short-sale constraints is always greater than the unconstrained price:

$$R_f(P_0^{con} - P_0^{unc}) = \gamma\sigma^2 + \Delta - 2\sqrt{\Delta\gamma\sigma^2} = (\sqrt{\gamma\sigma^2} - \sqrt{\Delta})^2 > 0$$

- We see that short-selling constraints indeed increase the equilibrium price if $\Delta > \gamma\sigma^2$ and that more dispersion in beliefs lead to higher prices $\frac{\partial P_0^{con}}{\partial\Delta} > 0$.

16 Speculative behavior in a dynamic setting (Harrison and Kreps (1978))

HK propose a model where agents have heterogeneous beliefs and where they can dynamically trade a stock and where in equilibrium they will display speculative behavior in the sense that they will be willing to pay more for a stock if they can retrade than the most optimistic investor would be willing to pay if she were obliged to hold on to the stock forever. In a sense they show that all agents are willing to pay more to exploit the option to resell to the most optimistic investor in any given state. So crucially for this effect to operate it has to be that agents are heterogeneous and that the ranking of investors' optimism may change across states.

16.1 An example

- two types of risk-neutral agents $a = \{1, 2\}$ who trade
- one stock which pays a dividend $d_t \in \{0, 1\}$ which follows a continuous time Markov chain.
- The transition probability Matrix perceived by agent a is Q^a with elements $q^a_{i,j}$, which denotes the probability of transitionning from state i to state j.
- for agent 1 we assume $q^1_{01} = 1/2 \;\; q^1_{10} = 2/3$
- for agent 2 we assume $q^2_{01} = 1/3 \;\; q^2_{10} = 1/4$
- There is one risk-free rate R_f so that the gross risk-free discount rate is $\gamma = 1/R_f = 0.75$.
- Short sales are not allowed.
- What it the value to agent a of holding the stock forever?

  Define p^a_i the value to a of holding the stock forever starting in state i. Clearly we have

  $$p^a_i = \gamma(q^a_{ii} p^a_i + q^a_{ij} p^a_j)$$

  which a system of two equations and 2 unknowns which can be solved for the two values for agent a. Using the numerical values we find:

  $$p^1_0 = 1.33 \;\; p^1_1 = 1.22 \;\text{ whereas }\; p^2_0 = 1.45 \;\; p^2_1 = 1.91$$

- Based on these values one might think that agent 2 should always hold the stock. However, HK shows that the prices p^2_i cannot be an equilibrium. Indeed, suppose that agent 1 anticipates these prices, then consider the strategy where she buys the stock in state 0 in order to sell it in state 1 at 1.91. Agent 1's expected value of this strategy in state 0 is then

  $$\begin{aligned} V_0 &= \gamma(q^1_{01}(1 + 1.91) + q^1_{00})(\gamma q^1_{01}(1 + 1.91) + q^1_{00}(\gamma q^1_{01}(1 + 1.91) + \ldots \\ &= \gamma(q^1_{01}(1 + 1.91) + q^1_{00} V_0) \end{aligned}$$

from which we obtain that

$$V_0 = 1.75 > 1.45$$

Thus agent 1 would be willing to bid up the price above p_0^2 to at least $1.75 > \max\{p_0^1, p_0^2\}$. But of course, then p_1^2 is too low a price as well, since agent 2 will now want to buy in state 1 to resell to agent 1 whenever state 0 comes around at what she will consider inflated prices $1.75 > 1.45 \ldots$

- What is then the equilibrium price in this market?

- HK define the set of consistent prices to satisfy

$$p_t(x_t) = \max_{a \in A} \sup_T E^a[\sum_{k=t+1}^{T} \gamma^{k-t} d_k(\xi_k) + \gamma^T p_T(\xi_T) | \xi_t = x_t] \ \ (\star)$$

where ξ_t is the (vector)-process that captures all of the information relevant to determine the dividend process $d_t(\xi_t)$ and A is the set of all agents who have different beliefs about the probability distribution of ξ_t and T is a stopping time defined with respect to the filtration generated by ξ_t. This is clearly the maximum value of any trading strategy that can be followed by any class of agent. If the price were lower, it would imply that there exists a class of agents a^* that has a trading strategy available (buy in state $\xi_t = x_t$ and sell according to the stopping time T^*) which would generate a higher valuation. All the other agents $a \neq a^*$ would like to sell, but because of the short-selling constraint they cannot.

- HK's proposition 1 simplifies the definition of consistent prices to

$$p_t(x_t) = \max_{a \in A} E^a[\gamma\{d_{t+1}(\xi_{t+1}) + p_{t+1}(\xi_{t+1})\} | \xi_t = x_t] \ \ (\star\star)$$

To prove the equivalence between the two statements, note that if $(\star\star)$ holds then $p_t(x_t) \geq E^a[\gamma\{d_{t+1}(\xi_{t+1}) + p_{t+1}(\xi_{t+1})\} | \xi_t = x_t]$. Thus using the law of iterated expectation and the optional stopping theorem we find that $(\star)$ holds as well. Conversely, it is clear that if $(\star)$ holds then clearly $p_t(x_t) \geq E^a[\gamma\{d_{t+1}(\xi_{t+1}) + p_{t+1}(\xi_{t+1}) | \xi_t = x_t]$. Now suppose that the strict inequality holds for some x_t then using the optional stopping theorem and law of iterated expectation we obtain that $p_t(x_t) > \max_{a\in A}\sup_T E^a[\sum_{k=t+1}^T \gamma^{k-t} d_k(\xi_k) + \gamma^T p_T(\xi_T)|\xi_t = x_t]$ which contradicts ($\star\star$).

- HK then give an algorithm to compute the optimal p_t^* following a recursive approach starting from $p_t^0 = 0$ then define the sequence p_t^n for all $n = 1, \ldots$ that satisfies:

$$p_t^n(x_t) = \max_{a\in A} E^a[\gamma\{d_{t+1}(\xi_{t+1}) + p_{t+1}^{n-1}(\xi_{t+1})|\xi_t = x_t]$$

  note that by definition the sequence is increasing so that the (possibly infinite) limit $\lim_{n\to\infty} p_t^n(x) = p_t^*(x)$ exists

- Proposition 2 then shows that p_t^* thus defined is (i) consistent in that it satisfies ($\star\star$) and (ii) minimal in that any other solution to ($\star\star$) will be greater or equal. The proof relies on taking the limit on both sides of the definition of p_t^n and using monotone convergence one sees that p_t^* clearly satisfies ($\star\star$). Then one can show recursively that any consistent pricing scheme $p_t(x_t) > 0 = p^0(x_t)$ (since we assume strictly positive dividends) and then we can show recursively that if a consistent price function $p_t(x_t) > p^n(x_t)$ then $p_t(x_t) > p^{n+1}(x_t)$ which implies that $p_t(x_t) > p^*(x_t)$.

- We note that since $T = \infty$ (i.e., buy and hold) is a feasible trading strategy, the consistent price must be larger than the maximum buy and hold value perceived by any agent.

- Going back to our example looking for a stationary solution p_0^*, p_1^* must satisfy:

$$p_0^* = \gamma \max_a\{q_{01}^a p_1^* + q_{00}^a p_0^*\}$$
$$p_1^* = \gamma \max_a\{q_{10}^a p_0^* + q_{11}^a p_1^*\}$$

  which gives: the solution $p_1^* = 2.07692$ and $p_0^* = 1.84615$. and we can compute the value for each agent in each state and find that

$$
\begin{aligned}
q_{01}^1 p_1^* + q_{00}^1 p_0^* &= p_0^* \\
q_{01}^2 p_1^* + q_{00}^2 p_0^* &= 1.69231 \\
q_{10}^1 p_0^* + q_{11}^a p_1^* &= 1.69231 \\
q_{10}^2 p_0^* + q_{11}^a p_1^* &= p_1^*
\end{aligned}
$$

Thus agent 1 holds the stock in state 0 whereas agent 2 holds the stock in state 1.

- HK's conclusion page 335 is worth reading. They point out that their analysis requires that agents have full knowledge of the other investors (different) beliefs and they agree to disagree, in order to derive the optimal trading strategy.

> *If one drops this utopian assumption, and further introduces such a real-life phenomenon as privileged information, one gets a world in which investors must turn to public information, such as prices and trading volume, to discover what their fellow investors know and how they will react to incoming information. At the risk of gross overstatement, we suggest that this line of reasoning might lead to a "legitimate" theory of technical analysis. Proponents of the efficient market hypothesis conclude that the rational portfolio strategy (in view of transaction costs and risk aversion) is to buy a well-diversified portfolio and hold it. Quite a different view of "rational" portfolio management emerges from our model. [...] In the general model, investors can achieve an expected net present value of zero only from stock bought in certain circumstances and only if they follow certain selling strategies. (The strategy of selling after one period, which leads to much churning of the portfolio, always works.) The strategy of buying in favorable circumstances and holding for many periods typically yields an expected loss. In brief, all investors must actively manage their portfolios in order to expect a proper return. [...]*
>
> *In our model, all investors have complete information from the outset, but still they arrive at different subjective assessments. Speculation and active portfolio management follow inevitably.*

17 Continuous time model: Scheinkman and Xiong (2003)

SX propose a continuous time version of the HK model, where disagreement is driven by overconfidence and biased learning. Their purpose is to study the joing dynamics of asset pricing bubbles, trading volume, and price volatility.

- The risky asset pays a dividend

$$dD_t = f_t dt + \sigma_D dZ_t^D$$
$$df_t = \lambda(\bar{f} - f_t)dt + \sigma_f dZ_t^f$$

- There are two sets of risk-neutral agents who each observe two signals about the fundamental driving the dividend process:

$$ds_t^j = f_t dt + \sigma^j dZ_t^j \;\; j = A, B$$

  Investors have to learn f_t from the history of D_t, s_t^A, s_t^B.

- Investors are over-confident in that the think their own signal s_t^j for group j is more precise than what it really is. So for example investors in group A think that:

$$ds_t^A = f_t dt + \sigma^A(\phi dZ_t^f + \sqrt{1-\phi^2} dZ_t^A)$$

  So they think their own signal has a correlation ϕ with the true fundamental, but that s_t^B is (correctly) uncorrelated with f_t so that the other investor has an uninformative signal.

- Based on this SX characterize the learning dynamics of each group of investor using standard Kalman filtering. So for example, for group A lets define $\hat{f}_t^A = E_t^A[f_t]$ and $\gamma_t = V_t^A[f_t]$ then we have

$$
\begin{aligned}
d\hat{f}_t^A &= \lambda(\bar{f} - \hat{f}_t^A)dt + \lambda_t^A(ds_t^A - \hat{f}_t^A dt) + \lambda_t^B(ds_t^B - \hat{f}_t^A dt) + \lambda_t^S(dD_t - \hat{f}_t^A dt) \\
\lambda_t^A &\approx \frac{Cov(f_t + df_t, ds_t^A)}{V[ds_t^A]} = \frac{\sigma_s \sigma_f \phi + \gamma_t}{\sigma_s^2} \\
\lambda_t^B &\approx \frac{Cov(f_t + df_t, ds_t^B)}{V[ds_t^B]} = \frac{\gamma_t}{\sigma_s^2} \\
\lambda_t^D &\approx \frac{Cov(f_t + df_t, dD_t)}{V[ds_t^B]} = \frac{\gamma_t}{\sigma_D^2}
\end{aligned}
$$

and where the posterior variance of the signal has the dynamics:

$$
d\gamma_t = (\sigma_f^2 - (\lambda_t^A)^2\sigma_s^2 - (\lambda_t^B)^2\sigma_s^2 - (\lambda_t^D)^2\sigma_D^2)dt
$$

HK focus on the stationary solution where the posterior variance is γ defined as the positive solution of the quadratic equation $d\gamma_t = 0$.

- HK then compute the dynamics of disagreement: the differences in beliefs process for group A: $g_t^A = \hat{f}_t^B - \hat{f}_t^A$ which they charaterize in proposition 1 as a simple mean-reverting process

$$
dg_t^A = \rho g_t^A dt + \sigma_g dW_g^A(t)
$$

This follows directly from the definition of $\hat{f}_t^A$ and the analogous expression for $\hat{f}_t^B$, taking their difference and using the expression for γ to simplify the expression for ρ.

- What is the value of the stock if agents A, B can trade with each other, are risk-neutral, and short-selling is not allowed? SX add the requirement that when an owner sells the asset she incurs trading costs c (otherwise agents might continuously sell the asset back and forth). Following HK (compare also with Amihud-Mendelson) they define the value of the stock to the owner $o = A, B$ as :

$$
p_t^o = \sup_{\tau} E^o[\int_t^{t+\tau} e^{-r(s-t)}dD_s + e^{-r\tau}(p_\tau^{\bar{o}} - c)]
$$

where$p_\tau^{\bar{o}}$ is defined as the reservation value of the current owner at the next transaction date (which will be the value to the other group).

we can rewrite this expression as:

$$p_t^o = \sup_\tau E^o[\int_t^\infty e^{-r(s-t)}dD_s - e^{-r\tau}\int_{t+\tau}^\infty e^{-r(s-t-\tau)}dD_s + e^{-r\tau}(p_\tau^{\bar{o}} - c)]$$

Using the definition of the dividend process it becomes:

$$p_t^o = \sup_\tau E^o[\frac{\bar{f}}{r} + \frac{\bar{f} - \hat{f}_t^o}{r+\lambda} - e^{-r\tau}(\frac{\bar{f}}{r} + \frac{\bar{f} - \hat{f}_\tau^o}{r+\lambda}) + e^{-r\tau}(p_\tau^{\bar{o}} - c)]$$

so it is natural to seek a solution of the form

$$p_t^o = \frac{\bar{f}}{r} + \frac{\bar{f} - \hat{f}_t^o}{r+\lambda} + q(g_t^o)$$

where the resale option value $q(g_t^o)$ will solve the American option pricing problem given by equation (13) in the paper, namely the value of the resale option satifies:

$$q(g_t^o) = \sup_\tau E^o[e^{-r\tau}(\frac{g_t^o}{r+\lambda} + q(g_\tau^{\bar{o}}) - c)]$$

- Theorem 2 then establishes that there exists an optimal barrier policy, so that when the differences in beliefs process becomes large enough $g_\tau^o > k^*$, then at that time τ it becomes optimal to pay the cost c and sell the asset to the second group of investors, at which point the process g_τ^o is reset for the new owners at $-k^*$ and the process restarts.

- The trading volume will depend directly on c (it goes to zero as c becomes very large and to infinity as it becomes very small and the switch between owners becomes continuous). SX show that even when $c \to \infty$ and the asset switches continuously between groups, the resale option value retains a strictly positive value.

- Figure 1 shows comparative statics of the model with respect to increases in the "overconfidence" parameter ϕ. Specifically, it shows the impact of overconfidence (which governs the cross-sectional dispersion in beliefs) on the time between trades (i.e, volume of trading), and the bubble component ($q(-k^*)$), and the excess volatility due to the bubble component (i.e., the volatilility of $q(g_t^0)$ evaluated at $g = k^*$).

- SX then consider different applications of their model, such as (i) how could it explain a market Krach (e.g., if there is a small probability that the fundamental value will be revealed to all), (ii) the impact of introducing a financial transactions' cost (e.g., raising c) on the bubble and volume dynamics, see figure 3; (iii) negative stub-values, and (iv) IPO underpricing.

  From their conclusion:

  > *This allows us to characterize properties of the magnitude of the bubble, trading frequency, and asset price volatility and to show that the model is consistent with the observation that in actual historical bubbles, volatility and turnover are also inordinate. Theoretical results and numerical exercises suggest that a small trading tax may be effective in reducing speculative trading, but it may not be very effective in reducing price volatility or the size of the bubble. Through a simple example, we also illustrate that the bubble can cause the price of a subsidiary to be larger than that of its parent firm, a violation of the law of one price. It is natural to conjecture that the existence of a speculative component in asset prices has implications for corporate strategies. Firm managers may be able to profit by adopting strategies that boost the speculative component. The underpricing of a firm's initial public offering (IPO) has been puzzling. Rajan and Servaes (1997) show that higher initial returns on an IPO lead to more analysts and media coverage. Since investors may disagree about the precision of information provided by the media, the increase in this coverage could increase the option component of the stock. Therefore, IPO underpricing could be a strategy used by firm managers to boost the price of their stocks.*