Kasper Larsen
Department of Mathematical Sciences
CARNEGIE MELLON UNIVERSITY

Market Microstructure (Kyle models)

General information: Market microstructure theory has been and continues to be an active reasearch area for both academics and practitioners (see e.g., the journals *Journal of Financial Markets* and *Market Microstructure and Liquidity*). While the formal definition of market mircrostructure may not be universally agreed upon (but see `https://en.wikipedia.org/wiki/Market_microstructure` for one such definition), it is clear that Kyle models and the many variants and extensions hereof play important roles in said theory. In these lectures we will try to cover the basics of Kyle's original model and its extension to multi agent settings.

We will make an effort to keep everything as simple as possible (e.g., we will only use finite discrete-time so we will not see any local martingales in these lectures...). The required knowledge to follow these lectures is basic conditional probability such as problems 1.1 and 1.2 below.

Lecture content:

1. One period Gaussian models: Kyle's model and Grossman and Stiglitz's model. We will follow (7) and (3).
2. Multi-period Kyle models: We will follow (7).
3. Multi-agent extensions of Kyle's model (see (6), (4), and (5)). We will follow (5).

Continuous-time papers:

We will not have time to discuss the following papers which constitute generalizations of the above mentioned papers (note that Kyle's original paper (7) contains the Gaussian continuous-time model):

- Continuous-time extensions of Kyle's model to the non-Gaussian case (see (1)).
- Continuous-time multi-agent extensions of Kyle's model (see (2)).

References

[1] Back, K., 1992: Insider Trading in Continuous Time, *Review of Financial Studies*, 5, 387–409.

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

[3] Grossman, S. J., and J. E. Stiglitz, 1980: On the Impossibility of Informationally Efficient Markets, *American Economic Review*, 70, 393-408.

[4] Foster, F., and S. Viswanathan, 1994: Strategic Trading with Asymmetrically Informed Traders and Long-Lived Information, *Journal of Financial and Quantitative Analysis*, 29, 499–518.

[5] Foster, F., and S. Viswanathan, 1996: Strategic Trading when Agents Forecast the Forecasts of Others, *Journal of Finance*, 51, 1437–1478.

[6] Holden, C.W., and A. Subrahmanyam, 1992: Long-Lived Private Information and Imperfect Competition, *Journal of Finance*, 47, 247–270.

[7] Kyle, A., 1985: Continuous Auctions and Insider Trading, *Econometrica*, 53, 1315–1336.

1 Exercises on Kyle's model

1. We are and will be using the projection theorem for Gaussian random variables: For zero-mean jointly normals $(\tilde{v}, \tilde{a})$ show

$$\mathbb{E}[\tilde{v}|\sigma(\tilde{a})] = \frac{\operatorname{cov}(\tilde{v}, \tilde{a})}{\mathbb{V}[\tilde{a}]}\tilde{a}.$$

$\diamondsuit$

2. Let A, B, and X be three random variables with $(A, X) \perp B$. Show

$$\mathbb{E}[X|\sigma(A, B)] = \mathbb{E}[X|\sigma(A)].$$

$\diamondsuit$

3. Let $\tilde{v}$ be any zero mean and square integrable random variable and let $(y_n)_{n\in\mathbb{N}}$ be any sequence of random variables (neither $\tilde{v}$ or $(y_n)_{n\in\mathbb{N}}$ are necessarily Gaussian). Explain why the function

$$\Sigma_n := \mathbb{E}\Big[\Big(\tilde{v} - \mathbb{E}[\tilde{v}|\sigma(y^n)]\Big)^2\Big], \quad n \in \mathbb{N},$$

is non-increasing. Give an example for which $\Sigma_\infty > 0$ and give a lemma with a set of sufficient conditions that ensure $\Sigma_\infty := \lim_{n\to\infty} \Sigma_n = 0$.

$\diamondsuit$

4. Implement the standard Kyle model for the parameters

$$N := 10, \quad \sigma_w := 2, \quad \sigma_{\tilde{v}} := 1.$$

Let us first check that we agree on the code. For these parameters I get the terminal variance to be $\Sigma_N = .11$.

$\diamondsuit$

5. Increase N and study numerically the limiting behavior of the value function coefficients $(\alpha_n)_{n=1}^N$. Do you see a facelift in the sense that as N gets bigger and bigger, α's slope over the interval $[1-\Delta, 1]$, i.e.,

$$\frac{\alpha_N - \alpha_{N-1}}{\Delta},$$

gets steeper and steeper?

$\diamondsuit$

6. One can of course view any equilibrium problem as a fixed point problem so let us try the following policy iteration algorithm: Fix the initial $\beta \in \mathbb{R}^N$ as the equilibrium strategy you found in a previous question. Compute (Σ, λ) by using Kyle's equations (3.18) and (3.19). Then update the coefficients (α, β) by using Kyle's equations (3.15) and (3.17). What happens numerically as you continue looping?

$\diamondsuit$

7. This question separates the forward component (filtering) from the backward component (optimization) in Kyle's standard model. We define

$$\gamma_n := \beta_n^2 \Sigma_n, \quad n = 1, ..., N-1, \quad \text{and} \quad \gamma_N := \frac{\sigma_w^2}{2\Delta}. \tag{1.1}$$

Use Kyle's equations (3.15)-(3.19) to derive the recursion

$$\gamma_{n+1} = \gamma_n \sigma^2 \frac{(\sigma_w^2 - \gamma_n \Delta)^2}{\gamma_n^3 \Delta^3 + 2\gamma_n^2 \Delta^2 \sigma_w^2 - 3\gamma_n \Delta \sigma_w^4 + \sigma_w^6}. \tag{1.2}$$

Similar to Kyle's original existence argument, this is a 3rd degree polynomial and we seek a root such that $2\gamma_n \Delta \in (0, \sigma_w^2)$. In that case, the second-order condition holds

$$1 > \frac{(\sigma_w^2 - 2\gamma_n \Delta)}{2(\sigma_w^2 - \gamma_n \Delta)} = \alpha_n \lambda_n. \tag{1.3}$$

Given the values γ_n, the forward component Σ_n is uniquely determined by Σ_0 and Kyle's equations (3.18) and (3.19).

Implement this idea numerically and check that you get the same results as you got previously.

$\diamondsuit$

8. In Kyle's 1985 model we replace the risk neutral insider by an insider with the exponential utility function (Constant Absolute Risk Aversion, CARA)

$$U(w) := -e^{-aw}, \quad w \in \mathbb{R},$$

where $a > 0$ is a constant. Set $\sigma_{\tilde{v}} := \sigma_w := 1$ and plot numerically as a function of $a > 0$ the two constants (β, λ) from this equilibrium together with the corresponding constants coming from the risk neutral insider equilibrium. How do you get the risk neutral equilibrium from the risk averse equilibrium?

$\diamondsuit$

2 Exercises on Grossman-Stiglitz's model

Let us consider the one period setting in Grossman-Stiglitz (1980) where (θ, x, ϵ) are independent Gaussian random variables with zero means and say unit variances. The interest rate is taken to be zero and we define the terminal stock value to be

$$\theta + \epsilon.$$

There are $M_I \in \mathbb{N}$ informed investors (who know θ and the price p) and $M_U \in \mathbb{N}$ uninformed investors (who only know the price p). They all have exponential utility functions given by $U(w) := -e^{-aw}$ for $w \in \mathbb{R}$ and $a > 0$. We assume that all initial holdings of the stock and money market account are zero.

1. *Radner*: Find equations for constants (α, β) and model parameters restrictions (if any) such that

   $$p = \alpha x + \beta\theta$$

   is an equilibrium price where $-x$ denotes the noise trades.

   $\diamondsuit$

2. *Nash*: (i) Conjecture the forms for the informed orders $A(p-\theta)$ and uninformed orders Bp for constants (A, B). For p implicitly defined by the market clearing condition

   $$x = M_I A(p-\theta) + (M_U - 1)Bp + x_U,$$

   solve the uninformed investor's problem (who uses the control x_U and only sees p). The condition that the optimal x_U should be Bp gives you one equation.

   (ii) Similarly, for p implicitly defined by the market clearing condition

   $$x = (M_I - 1)A(p-\theta) + M_U Bp + x_I,$$

   solve the informed investor's problem (who uses the control x_I and sees both p and θ). The condition that the optimal x_I should be $A(p-\theta)$ gives you a second equation.

   $\diamondsuit$

3. Find model parameter restrictions under which these two equations produce an equilibrium described by (A, B) (don't forget the second order conditions). Compare the Radner and Nash equilibria whenever they both exist.

$\diamondsuit$

3 Exercises on discrete-time multi-agent Kyle extensions

1. Implement the model introduced in Foster and Viswanathan (1996) numerically for the parameters

$$I := 2, \quad \sigma_w := 2, \quad \rho := 0.25, \quad \sigma_{\tilde{a}} := 1, \quad N := 10. \tag{3.1}$$

Plot the trajectories of λ_n, β_n, and α_n for the parameter values (3.1).

$\diamondsuit$

2. An important feature of Foster and Viswanathan (1996) is the emergence of negative correlation, i.e.,

$$\rho_n := \mathbb{E}[(\tilde{a}^i - \hat{t}_n)(\tilde{a}^j - \hat{t}_n)], \quad i \neq j,$$

can be negative even though $\mathrm{cov}(\tilde{a}^i, \tilde{a}^j) > 0$. In the setting of Foster and Viswanathan (1996), prove the following:

$$\Delta\rho_n = \Delta\Sigma_{1,n}, \quad \Sigma_{1,n} := \mathbb{E}[(\tilde{a}^i - \hat{t}_n)^2].$$

Plot the trajectory of ρ_n as well as $\mathrm{corr}(\Delta\theta_n^1, \Delta\theta_n^2)$ for the parameter values (3.1). Here $(\Delta\theta_n^1, \Delta\theta_n^2)$ denote the optimal strategies in equilibrium.

$\diamondsuit$

3. If you don't know this result already, you need to prove it: Let (A, B) be jointly normally distributed with correlation ρ. Show that $A \perp B$ if and only if $\rho = 0$.

   Use this to show that in Foster and Viswanathan (1996) we have $\Delta w_{i,n} \perp \sigma(\tilde{a}^i, w_i^{n-1})$. Subsequently, show that $w_{i,n} = \sum_{k=1}^{n} \Delta w_{i,k} \perp \tilde{a}^i$.

   ♢

4. (Nested information and asymmetric information) This model is from Foster and Viswanathan (1994). We consider a Kyle setting with two informed investors: Agent I knows $(\tilde{v}, \tilde{a})$ whereas Agent L only knows $\tilde{a}$. Perform the same analysis as we did for Foster and Viswanathan (1996). To get started, consider strategies for the two investors of the forms:

$$\Delta\theta_n^I = \beta_n^I(\tilde{v} - s_{n-1}) + \alpha_n^I(s_{n-1} - p_{n-1}), \tag{3.2}$$
$$\Delta\theta_n^L = \beta_n^L(s_{n-1} - p_{n-1}), \tag{3.3}$$

   where in equilibrium we wish to have

$$p_n = \mathbb{E}[\tilde{v}|\sigma(y^n)], \quad s_n = \mathbb{E}[\tilde{v}|\sigma(\tilde{a}, y^n)]. \tag{3.4}$$

   ♢

5. (Not nested and asymmetric information) This model is from Appendix C in Choi, Larsen, and Seppi (2016) and is a variant of Foster and Viswanathan (1994) we considered in the previous question. We consider a Kyle setting with two informed investors: Agent I knows $\tilde{v}$ whereas Agent L knows $\tilde{a}$. Adjust your analysis in the previous question to this case.

   Hint: Because Agent I cannot observe the process s, Agent I cannot use the strategy (3.2) *unless $\beta^I = \alpha^I$.*

   ♢

4 Notes on Foster and Viswanathan (1996)

We consider $\tilde{v} := \sum_{i=1}^{I} \tilde{a}^i$ where $\tilde{a}^1, ..., \tilde{a}^I$ are zero-mean jointly normals with identical variances and covariances. We assume that $\tilde{a}^1, ..., \tilde{a}^I$ are independent of the Brownian increments $\Delta w_n \sim \mathcal{N}(0, \sigma_w^2 \Delta)$. We start by defining the Gaussian hat-processes:

$$\Delta\hat{\theta}_n^i := \beta_n(\tilde{a}^i - \hat{t}_{n-1}), \quad \hat{\theta}_0^i := 0, \tag{4.1}$$

$$\Delta\hat{y}_n := \sum_{i=1}^{I} \Delta\hat{\theta}_n^i + \Delta w_n, \quad \hat{y}_0 := 0, \tag{4.2}$$

$$\Delta\hat{p}_n := \lambda_n \Delta\hat{y}_n, \quad \hat{p}_0 := 0, \tag{4.3}$$

$$\Delta\hat{t}_n := \zeta_n \Delta\hat{y}_n, \quad \hat{t}_0 := 0. \tag{4.4}$$

Step 1: We wish to have the relations (with $\hat{t}$ independent of i):

$$\hat{p}_n = \mathbb{E}[\tilde{v}|\sigma(\hat{y}^n)] = \sum_j \mathbb{E}[\tilde{a}^i|\sigma(\hat{y}^n)] = I\hat{t}_n. \tag{4.5}$$

We define the unconditional expectations

$$\Sigma_{1,n} := \mathbb{E}[(\tilde{a}^i - \hat{t}_n)^2], \quad \Sigma_{1,0} = \sigma_{\tilde{a}}^2, \tag{4.6}$$

$$\Sigma_{2,n} := \mathbb{E}[(\tilde{a}^i - \hat{t}_n)\sum_{j=1}^{I}(\tilde{a}^j - \hat{t}_n)], \quad \Sigma_{2,0} = \sigma_{\tilde{a}}^2 + (I-1)\text{cov}(\tilde{a}^i, \tilde{a}^j), \tag{4.7}$$

$$\Sigma_{3,n} := \mathbb{E}[(\tilde{v} - \hat{p}_n)(\tilde{a}^i - \hat{t}_n)]. \tag{4.8}$$

As we shall see everything can be expressed in terms of only one of these three quantities. Indeed, by iterated expectations, we have

$$\begin{aligned}
\Sigma_{3,n} &= \mathbb{E}[(\sum_j \tilde{a}^j - \hat{p}_n)(\tilde{a}^i - \hat{t}_n)] \\
&= \sum_j \mathbb{E}[\tilde{a}^j(\tilde{a}^i - \hat{t}_n)] \\
&= \sum_j \mathbb{E}[(\tilde{a}^j - \hat{t}_n)(\tilde{a}^i - \hat{t}_n)] \\
&= \Sigma_{2,n}.
\end{aligned}$$

The ith investor's innovation process is defined by $w_{i,0} := 0$ and

$$\begin{aligned}
\Delta w_{i,n} &:= \beta_n \sum_j \Big(\tilde{a}^j - \hat{t}_{n-1} - \mathbb{E}[\tilde{a}^j - \hat{t}_{n-1} | \sigma(\tilde{a}^i, \hat{y}^{n-1})] \Big) + \Delta w_n \\
&= \Delta \hat{y}_n - \beta_n \sum_j \mathbb{E}[\tilde{a}^j - \hat{t}_{n-1} | \sigma(\tilde{a}^i, \hat{y}^{n-1})] \\
&= \Delta \hat{y}_n - \beta_n \frac{\Sigma_{2,n-1}}{\Sigma_{1,n-1}} (\tilde{a}^i - \hat{t}_{n-1}).
\end{aligned}$$

This should be seen in contrast to Kyle (1985) where $\Delta w = \Delta \hat{y}_n - \Delta \theta_n$ is the insider's innovation process. The market makers' innovation process is defined by $w_{M,0} := 0$ and

$$\begin{aligned}
\Delta w_{M,n} &:= \beta_n \sum_j (\tilde{a}^j - \hat{t}_{n-1}) + \Delta w_n \\
&= \Delta \hat{y}_n.
\end{aligned}$$

There are a couple of things to note about the innovation processes:

- They are zero mean Gaussian processes.
- We have $\sigma(\tilde{a}^i, \hat{y}^n) = \sigma(\tilde{a}^i, w_i^n)$ and $\sigma(\hat{y}^n) = \sigma(w_M^n)$.
- $\Delta w_{i,n} \perp \sigma(\tilde{a}^i, \hat{y}^{n-1})$ because for $f \in \{\tilde{a}^i, \hat{y}_1, ..., \hat{y}_{n-1}\}$ we have

  $$\mathbb{E}[\Delta w_{i,n} f] = \mathbb{E}\Big[\mathbb{E}[\Delta w_{i,n} f | \sigma(\tilde{a}^i, \hat{y}^{n-1})]\Big] = \mathbb{E}\Big[f \mathbb{E}[\Delta w_{i,n} | \sigma(\tilde{a}^i, \hat{y}^{n-1})]\Big].$$

  The claim then follows from the joint normality.

We will next compute the pricing coefficients and we start by noticing:

$$\begin{aligned}
\mathbb{V}[\Delta w_{M,n}] &= \beta_n^2 \mathbb{V}\left[\sum_j (\tilde{a}^j - \hat{t}_{n-1}) \right] + \sigma_w^2 \Delta \\
&= \beta_n^2 I \Sigma_{2,n-1} + \sigma_w^2 \Delta.
\end{aligned}$$

The dynamics of the price process can then be found as

$$\begin{aligned}
\Delta\hat{p}_n &= \mathbb{E}[\tilde{v}|\sigma(\hat{y}^n)] - \hat{p}_{n-1} \\
&= \mathbb{E}[\tilde{v} - \hat{p}_{n-1}|\sigma(\Delta w_{M,n})] \\
&= \frac{\mathbb{E}[(\tilde{v} - \hat{p}_{n-1})\Delta w_{M,n}]}{\mathbb{V}[\Delta w_{M,n}]}\Delta w_{M,n}.
\end{aligned}$$

Because $\Delta w_{M,n} = \Delta\hat{y}_n$ this produces the requirement

$$\begin{aligned}
\lambda_n &= \frac{\mathbb{E}[(\tilde{v} - \hat{p}_{n-1})\Delta w_{M,n}]}{\mathbb{V}[\Delta w_{M,n}]} \\
&= \frac{\beta_n \mathbb{E}[(\tilde{v} - \hat{p}_{n-1})\sum_j (\tilde{a}^j - \hat{t}_{n-1})]}{\beta_n^2 I\Sigma_{2,n-1} + \sigma_w^2\Delta} \\
&= \frac{\beta_n I\Sigma_{3,n-1}}{\beta_n^2 I\Sigma_{2,n-1} + \sigma_w^2\Delta}.
\end{aligned}$$

In a similar manner we find the representation

$$\Delta\hat{t}_n = \frac{\mathbb{E}[(\tilde{a}^i - \hat{t}_{n-1})\Delta w_{M,n}]}{\mathbb{V}[\Delta w_{M,n}]}\Delta w_{M,n}.$$

So we need

$$\zeta_n = \frac{\beta_n\Sigma_{2,n-1}}{\beta_n^2 I\Sigma_{2,n-1} + \sigma_w^2\Delta}.$$

Because $\Sigma_2 = \Sigma_3$ we must have

$$\frac{\lambda_n}{\zeta_n} = I.$$

Finally, we need the filter dynamics

$$\begin{aligned}
\Sigma_{1,n} &= \mathbb{V}[\tilde{a}^i - \hat{t}_{n-1} - \Delta\hat{t}_n] \\
&= \mathbb{V}[\tilde{a}^i - \hat{t}_{n-1} - \zeta_n\Delta\hat{y}_n] \\
&= \mathbb{V}[\tilde{a}^i - \hat{t}_{n-1} - \zeta_n\beta_n\sum_j(\tilde{a}^j - \hat{t}_{n-1}) - \zeta_n\Delta w_n] \\
&= \Sigma_{1,n-1} + \zeta_n^2\beta_n^2 I\Sigma_{2,n-1} - 2\zeta_n\beta_n\Sigma_{2,n-1} + \zeta_n^2\sigma_w^2\Delta, \\
\Sigma_{2,n} &= \mathbb{E}\Big[\Big(\tilde{a}^i - \hat{t}_{n-1} - \zeta_n\beta_n\sum_j(\tilde{a}^j - \hat{t}_{n-1}) - \zeta_n\Delta w_n\Big) \\
&\qquad \times \sum_k\Big(\tilde{a}^k - \hat{t}_{n-1} - \zeta_n\beta_n\sum_l(\tilde{a}^l - \hat{t}_{n-1}) - \zeta_n\Delta w_n\Big)\Big] \\
&= \mathbb{E}\Big[\Big(\tilde{a}^i - \hat{t}_{n-1} - \zeta_n\beta_n\sum_j(\tilde{a}^j - \hat{t}_{n-1}) - \zeta_n\Delta w_n\Big) \\
&\qquad \times \Big(\sum_k(\tilde{a}^k - \hat{t}_{n-1}) - \zeta_n I\beta_n\sum_l(\tilde{a}^l - \hat{t}_{n-1}) - \zeta_n I\Delta w_n\Big)\Big] \\
&= \Sigma_{2,n-1} - \zeta_n I\beta_n\Sigma_{2,n-1} - \zeta_n I\beta_n\Sigma_{2,n-1} + \zeta_n^2 I^2\beta_n^2\Sigma_{2,n-1} + I\zeta_n^2\sigma_w^2\Delta.
\end{aligned}$$

From the latter expression and the expression for ζ_n we get

$$\begin{aligned}
\Sigma_{2,n} &= \frac{\Sigma_{2,n-1}\sigma_w^2\Delta}{\sigma_w^2\Delta + \Sigma_{2,n-1}I\beta_n^2}, \\
\Sigma_{2,n-1} &= \frac{\Sigma_{2,n}\sigma_w^2\Delta}{\sigma_w^2\Delta - \Sigma_{2,n}I\beta_n^2}, \\
\Delta\Sigma_{2,n} &= -I\beta_n\zeta_n\Sigma_{2,n-1}, \\
\zeta_n &= \frac{\beta_n\Sigma_{2,n}}{\sigma_w^2\Delta}.
\end{aligned}$$

We also note that $\Delta\Sigma_2 = I\Delta\Sigma_1$. This property leads to the following two facts. First, we have the representation

$$\begin{aligned}
\Sigma_{2,n} &= \Sigma_{2,0} + \sum_k \Delta\Sigma_{2,k} \\
&= \Sigma_{2,0} + I\sum_k \Delta\Sigma_{1,k} \\
&= \sigma_{\tilde{a}}^2 + (I-1)\rho_0 + I(\Sigma_{1,n} - \sigma_{\tilde{a}}^2).
\end{aligned}$$

Second, we define the covariance function

$$\begin{aligned}\rho_n &:= \mathbb{E}[(\tilde{a}^i - \hat{t}_n)(\tilde{a}^j - \hat{t}_n)] \\ &= \frac{\Sigma_{2,n} - \Sigma_{1,n}}{I-1}.\end{aligned}$$

Then we have $\Delta\rho_n = \Delta\Sigma_{1,n}$ because

$$\begin{aligned}(I-1)\Delta\rho_n &= \Delta\Sigma_{2,n} - \Delta\Sigma_{1,n} \\ &= (I-1)\Delta\Sigma_{1,n}.\end{aligned}$$

This observation produces a main feature of the model: Even though $\rho_0 > 0$ we can have $\rho_n < 0$ provided that there are enough time steps.

Step 2: Next, we turn to the optimization problems and we define

$$\Delta p_n := \lambda_n \Delta y_n, \quad p_0 := 0, \tag{4.9}$$

$$\Delta t_n := \zeta_n \Delta y_n, \quad t_0 := 0. \tag{4.10}$$

Because $p_0 = \hat{p}_0 = 0$, $t_0 = \hat{t}_0 = 0$, and

$$\Delta\hat{p}_n - \Delta p_n = \lambda_n \sum_j (\Delta\hat{\theta}^j_n - \Delta\theta^j_n) = I(\Delta\hat{t}_n - \Delta t_n),$$

we have

$$\hat{p}_n - p_n = I(\hat{t}_n - t_n).$$

The ith investor seeks to maximize

$$\mathbb{E}\left[\sum_{n=1}^{N} (\tilde{v} - p_n)\Delta\theta^i_n \Big| \sigma(\tilde{a}^i)\right], \tag{4.11}$$

over controls $\Delta\theta^i_n \in \sigma(\tilde{a}^i, y^{n-1})$. We fix i and we fix the functional form of agent j's strategy to be

$$\Delta\theta^j_n = \beta_n(\tilde{a}^j - t_{n-1}), \quad j \neq i. \tag{4.12}$$

A couple of things to note:

- The realization of θ^j, $j \neq i$, is not fixed because it depends on the realization of t_n which in turn depends on θ^i.
- We never use hat-processes when fixing the other agents' strategies. For example, we will never use strategies like

  $$\beta_n(\tilde{a}^j - \hat{t}_{n-1}),$$

  for agent j with $j \neq i$. If we did that, the realization of Agent j's control would be independent of Agent i's choice of $\Delta\theta^i$.
- We will show next that the optimal strategy for investor i has the form

  $$\beta_n(\tilde{a}^i - \hat{t}_{n-1}) + \alpha_n(\hat{t}_{n-1} - t_{n-1}).$$

  However, we will never include off equilibrium deviation terms like $\alpha_n(\hat{t}_{n-1} - t_{n-1})$ for agent j when fixing agent j's strategy ($j \neq i$).

To move one, we observe

$$\sigma(\tilde{a}^i, y^n) = \sigma(\tilde{a}^i, \hat{y}^n) = \sigma(\tilde{a}^i, w_i^n).$$

The latter equality follows by the definition of the innovation process w_i. We will show the first equality by induction. The basis step is true because

$$\begin{aligned}
\sigma(\tilde{a}^i, y_1) &= \sigma(\tilde{a}^i, \beta_1 \sum_{j \neq i} \tilde{a}^j + \Delta\theta_1^i + \Delta w_1) \\
&= \sigma(\tilde{a}^i, \beta_1 \sum_{j \neq i} \tilde{a}^j + \Delta w_1) \\
&= \sigma(\tilde{a}^i, \hat{y}_1).
\end{aligned}$$

The next step follows similarly:

$$\begin{aligned}
\sigma(\tilde{a}^i, \hat{y}_1, \hat{y}_2) &= \sigma(\tilde{a}^i, \Delta\hat{y}_1, \Delta\hat{y}_2) \\
&= \sigma(\tilde{a}^i, \beta_1 \sum_{j} \tilde{a}^j + \Delta w_1, \beta_2 \sum_{j} (\tilde{a}^j - \hat{t}_1) + \Delta w_2) \\
&= \sigma(\tilde{a}^i, \beta_1 \sum_{j \neq i} \tilde{a}^j + \Delta w_1, \beta_2 \sum_{j \neq i} \tilde{a}^j + \Delta w_2).
\end{aligned}$$

On the other hand, we have

$$
\begin{aligned}
\sigma(\tilde{a}^i, y_1, y_2) &= \sigma(\tilde{a}^i, \Delta y_1, \Delta y_2) \\
&= \sigma(\tilde{a}^i, \beta_1 \sum_{j\neq i} \tilde{a}^j + \Delta\theta_1^i + \Delta w_1, \beta_2 \sum_{j\neq i} (\tilde{a}^j - t_1) + \Delta\theta_2^i + \Delta w_2) \\
&= \sigma(\tilde{a}^i, \beta_1 \sum_{j\neq i} \tilde{a}^j + \Delta w_1, \beta_2 \sum_{j\neq i} (\tilde{a}^j - t_1) + \Delta w_2).
\end{aligned}
$$

The next steps are identical.

For the optimization problem we first note

$$
\begin{aligned}
&\mathbb{E}[\Delta\hat{p}_n - \Delta p_n | \sigma(\tilde{a}^i, w_i^{n-1})] \\
&= \lambda_n \mathbb{E}\left[\sum_{j=1}^{I} \Delta\hat{\theta}_n^j - \sum_{j\neq i} \Delta\theta_n^j - \Delta\theta_n^i \Big| \sigma(\tilde{a}^i, w_i^{n-1}) \right] \\
&= \lambda_n \mathbb{E}\left[\sum_{j=1}^{I} \beta_n(\tilde{a}^j - \hat{t}_{n-1}) - \sum_{j\neq i} \beta_n(\tilde{a}^j - t_{n-1}) \Big| \sigma(\tilde{a}^i, w_i^{n-1}) \right] - \lambda_n \Delta\theta_n^i \\
&= \lambda_n \Big(\beta_n(\tilde{a}^i - \hat{t}_{n-1}) + \beta_n(I-1)(t_{n-1} - \hat{t}_{n-1}) - \Delta\theta_n^i \Big).
\end{aligned}
$$

From the definition of the ith investor's innovation process we get the representation

$$
\begin{aligned}
\Delta\hat{p}_n &= \lambda_n \Delta\hat{y}_n \\
&= \lambda_n \Big(\Delta w_{i,n} + \beta_n \frac{\Sigma_{2,n-1}}{\Sigma_{1,n-1}} (\tilde{a}^i - \hat{t}_{n-1}) \Big).
\end{aligned}
$$

Based on these two properties we find

$$
\begin{aligned}
&\mathbb{E}[(\tilde{v} - p_n)\Delta\theta_n^i | \sigma(\tilde{a}^i, w_i^{n-1})] \\
&= \Delta\theta_n^i \mathbb{E}[(\tilde{v} - \hat{p}_{n-1} + \hat{p}_{n-1} - p_{n-1} - \Delta\hat{p}_n + \Delta\hat{p}_n - \Delta p_n) | \sigma(\tilde{a}^i - \hat{t}_{n-1})] \\
&= \Delta\theta_n^i \Big(\frac{\Sigma_{3,n-1}}{\Sigma_{1,n-1}} (\tilde{a}^i - \hat{t}_{n-1}) + I(\hat{t}_{n-1} - t_{n-1}) - \lambda_n \beta_n \frac{\Sigma_{2,n-1}}{\Sigma_{1,n-1}} (\tilde{a}^i - \hat{t}_{n-1}) \\
&\qquad + \lambda_n \Big(\beta_n(\tilde{a}^i - \hat{t}_{n-1}) + \beta_n(I-1)(t_{n-1} - \hat{t}_{n-1}) - \Delta\theta_n^i \Big) \Big) \\
&= \Delta\theta_n^i \Big(\frac{\Sigma_{2,n-1}}{\Sigma_{1,n-1}} (1 - \lambda_n\beta_n) + \lambda_n\beta_n \Big) Y_{1,n-1} + \Delta\theta_n^i \big(I - \lambda_n\beta_n(I-1) \big) Y_{2,n-1} - \lambda_n (\Delta\theta_n^i)^2,
\end{aligned}
$$

where we have defined the two state processes

$$Y_{1,n} := \tilde{a}^i - \hat{t}_n, \quad Y_{2,n} := \hat{t}_n - t_n.$$

The state processes have the following Markovian dynamics

$$\begin{aligned}
\Delta Y_{1,n} &= -\Delta \hat{t}_n \\
&= -\zeta_n \Delta \hat{y}_n \\
&= -\zeta_n \Big(\Delta w_{i,n} + \beta_n \frac{\Sigma_{2,n-1}}{\Sigma_{1,n-1}} Y_{1,n-1} \Big), \\
\Delta Y_{2,n} &= \Delta \hat{t}_n - \Delta t_n \\
&= \zeta_n (\Delta \hat{y}_n - \Delta y_n) \\
&= \zeta_n \Big(\beta_n (\tilde{a}^i - \hat{t}_{n-1}) + \beta_n (I-1)(t_{n-1} - \hat{t}_{n-1}) - \Delta \theta_n^i \Big) \\
&= \zeta_n \Big(\beta_n Y_{1,n-1} - \beta_n (I-1) Y_{2,n-1} - \Delta \theta_n^i \Big).
\end{aligned}$$

We then conjecture the value function form

$$\sup_{\Delta\theta_k \in \sigma(\tilde{a}^i, y^{k-1})} \mathbb{E}[\sum_{k=n+1}^{N} (\tilde{v} - p_k)\Delta\theta_k | \sigma(\tilde{a}^i, y^n)] = I_n^{(0)} + I_n^{(1,1)} Y_{1,n}^2 + I_n^{(1,2)} Y_{1,n} Y_{2,n} + I_n^{(2,2)} Y_{2,n}^2.$$

Finally, to use dynamical programming, we need to compute the terms

$$\mathbb{E}[Y_{1,n}^2 | \sigma(\tilde{a}^i, w_i^{n-1})], \quad \mathbb{E}[Y_{1,n} Y_{2,n} | \sigma(\tilde{a}^i, w_i^{n-1})], \quad \mathbb{E}[Y_{2,n}^2 | \sigma(\tilde{a}^i, w_i^{n-1})].$$