15
The Kyle Model

Economists think about what people ought to do. Psychologists watch what they actually do.

(Daniel Kahneman)

As we discussed in Section 11.1, price impact can be interpreted in two different ways: as a statistical effect caused by a local order-flow imbalance in a market otherwise in equilibrium, or as the process by which revealed information is incorporated into the price. In the previous chapters, we have mainly focused on the first of these interpretations. In the present chapter, we turn our attention to the second interpretation. Specifically, we introduce the classic **Kyle model**, which seeks to shed light on the mechanism by which private information is gradually incorporated into prices, via trading. Kyle's original paper is often cited as the foundation of the field of market microstructure. The Kyle model shows how price impact can arise when market-makers anticipate adverse selection in an auction setting. As we discuss throughout the chapter, the model also makes clear why even informed traders should act cautiously and use small orders to ensure that they only reveal their information very slowly.

15.1 Model Set-Up

Consider a single-asset market populated by an **informed trader** (Alice), a **market-maker** (Bob), and a set of **noise traders**, who behave as follows:

- **Alice**: At time $t = 0$, Alice discovers some private information about the price p_F that the asset will have at $t = 1$. For example, Alice may be an insider trader who knows that at time $t = 1$, there will be a public announcement about a takeover bid at price p_F. Alice is the only market participant with access to this (private) information, so that Alice does not have to worry about other market participants trying to exploit the same information. Based on her private information, Alice chooses a volume Q of the asset to buy ($\varepsilon = +1$) or sell ($\varepsilon = -1$), in such a way as to maximise her expected profit, when discounting the expected impact cost of her own trade. Alice has no risk constraints that would limit her position.

- **Noise traders**: These uninformed traders do not have access to private information, but rather simply trade for idiosyncratic reasons (for a full discussion of the role of noise traders in the market ecology, see Chapter 1). In doing so, the noise traders generate a random order flow with a net volume of V_{noise}, whose sign and amplitude is independent from p_{F}.
- **Bob**: Bob is a market-maker. He clears the market by matching the net total buy or sell volume $\Delta V = V_{\text{noise}} + \varepsilon Q$ with his own inventory, at a clearing price $\widehat{p}$. Bob's choice of $\widehat{p}$ is rule-based, and such that he breaks even on average, in a sense that we make precise below. Bob has no inventory constraints that would limit his position.

At time $t = 1$, the price p_{F} is revealed. At this time, Alice's asset is exactly worth p_{F}. This assumes that she can buy or sell any quantity of the asset at price p_{F} without causing impact.

In the tradition of theoretical economics, one then looks for an equilibrium between Alice and Bob, such that:

(i) **Profit maximisation**: given Bob's price-clearing policy, Alice's signed volume εQ must maximise her expected gain. She buys/sells at the clearing price $\widehat{p}$, but the asset will in fact be worth p_{F}, so her gain is given by

$$\mathcal{G} = \varepsilon Q(p_{\text{F}} - \widehat{p}). \tag{15.1}$$

Note again that there is no risk constraint that would limit Alice's position in the asset.

(ii) **Market efficiency**: Bob's clearing price must be such that $\mathbb{E}_{\text{Bob}}[p_{\text{F}}|\Delta V] = \widehat{p}$, where the expectation is taken using the information available to Bob only (i.e. without the information known to Alice). This corresponds to a situation where Bob breaks even on average, given an incoming volume ΔV.

Given that Alice knows all of the above (i.e. she knows how both the noise traders and Bob will act in a given situation), how should she choose Q to maximise her expected profit $\mathbb{E}[\mathcal{G}]$ at time $t = 1$? For Alice to choose an optimal value of Q to maximise $\mathcal{G}$, she must consider Bob's clearing price. Specifically, Alice knows that Bob has a mechanical rule for choosing $\widehat{p}$ as a function of ΔV, so she must use this knowledge when deciding how to act.

Bob observes the net volume ΔV, but does not know the value of p_{F}. Since orders are anonymous, he also does not know the values of ε or Q. However, Bob knows that Alice knows his price-clearing rule, and also knows that she acts so as to optimise her gain with the information at her disposal (which he will try to infer from the knowledge of ΔV).

We further assume that Bob knows that the net volume executed by the noise traders is a Gaussian random variable with zero mean and standard deviation equal to Σ_V, and that the mispricing $p_F - p_0$ is similarly a Gaussian random variable with zero mean (over time) and standard deviation σ_F. The latter quantity is related to the typical amount of information available to insiders at time 0 but not yet included in the price.

15.2 Linear Impact

What are Alice's and Bob's optimal actions in this market? Consider the case where all random variables are Gaussian and where Bob's price-fixing rule is linear in the order imbalance, such that

$$\widehat{p} = p_0 + \Lambda \Delta V, \tag{15.2}$$

for some impact parameter Λ called **Kyle's lambda**. In this framework, one can show that the solution to this problem (or "market equilibrium") is self-consistent and can be fully determined. Indeed, profit maximisation on Alice's part leads to

$$\widehat{Q} = \mathrm{argmax}_Q \mathbb{E}[\mathcal{G}]; \qquad \mathcal{G} = \varepsilon Q \times (p_F - p_0 - \Lambda \Delta V). \tag{15.3}$$

Since the expected value of the random imbalance V_{noise} is zero, one has $\mathbb{E}[\Delta V] = \varepsilon Q$. This leads to a quadratic maximisation problem, with the solution

$$\widehat{Q} = \frac{1}{2} \frac{(p_F - p_0)}{\Lambda}. \tag{15.4}$$

This result means that Alice should choose $\widehat{Q}$ to be proportional to the mispricing $p_F - \widehat{p}$. Since Bob knows that Alice will do this, he attempts to estimate the value of $\varepsilon\widehat{Q}$, when he only observes ΔV. Estimating $\varepsilon\widehat{Q}$ allows him, using Equation (15.4), to guess the value of p_F used by Alice, and choose Λ in such a way that his clearing price $\widehat{p}$ is an unbiased estimate of the future price.

Using **Bayes' theorem**, the conditional probability that Alice's volume is $\varepsilon\widehat{Q}$, given ΔV, is

$$\begin{aligned} \mathbb{P}_{\text{Bob}}(\varepsilon\widehat{Q} \mid \Delta V) &\propto \mathbb{P}(\Delta V \mid \varepsilon\widehat{Q}) \times \mathbb{P}(\varepsilon\widehat{Q}), \\ &\propto \exp\left[-\frac{(\Delta V - \varepsilon\widehat{Q})^2}{2\Sigma_V^2}\right] \times \exp\left[-\frac{2(\varepsilon\widehat{Q})^2\Lambda^2}{\sigma_F^2}\right], \end{aligned}$$

where in the second exponential we have used the relation between $\widehat{Q}$ and $p_F - p_0$ from Equation (15.4). By merging the two exponentials together, we see that Bob's inferred distribution for $\varepsilon\widehat{Q}$ is Gaussian, with the following conditional mean:

$$\mathbb{E}_{\text{Bob}}[\varepsilon\widehat{Q} \mid \Delta V] = \sigma_F^2 \frac{\Delta V}{\sigma_F^2 + 4\Lambda^2\Sigma_V^2}. \tag{15.5}$$

 https://doi.org/10.1017/9781316659335.023

This in turn converts into Bob's best estimate of Alice's view on the future price, again using Equation (15.4)

$$\widehat{p} := \mathbb{E}_{\text{Bob}}\left[p_{\text{F}}|\Delta V\right] = p_0 + 2\Lambda\sigma_{\text{F}}^2 \frac{\Delta V}{\sigma_{\text{F}}^2 + 4\Lambda^2\Sigma_V^2}. \tag{15.6}$$

Identifying this expression with Equation (15.2), it also follows that

$$\Lambda = \frac{2\Lambda\sigma_{\text{F}}^2}{\sigma_{\text{F}}^2 + 4\Lambda^2\Sigma_V^2}.$$

Simplifying this expression provides the following solution for Kyle's lambda:

$$\Lambda = \frac{\sigma_{\text{F}}}{2\Sigma_V}. \tag{15.7}$$

By choosing this value of Λ, Bob thus makes sure that the strategic behaviour of Alice and the stochastic nature of the noise traders combine in such a way that the realised price $\widehat{p}$ is an unbiased estimate of the fundamental price p_{F}, given the publicly available information at time $t = 0$.

15.3 Discussion

The Kyle model raises several interesting points for discussion. First, the mean impact of an order grows linearly with Q: the Kyle model leads to a **linear impact law**. The linear scale factor Λ grows with the typical amount of private information present in the market (measured by σ_{F}) but decreases with the typical volume of uninformed trades (measured by Σ_V). This captures the basic intuition that market-makers protect themselves against adverse selection from informed traders by increasing the cost of trading, and benefit from the presence of noise traders to reduce price impact.

Second, using the result for $\widehat{Q}$ in Equation (15.4), the result for $\widehat{p}$ in Equation (15.6), and Equation (15.7), it follows that Alice's gain is given by

$$\begin{aligned}\mathbb{E}[\varepsilon\widehat{Q}\cdot(p_{\text{F}} - \widehat{p})] &= \frac{\sigma_{\text{F}}^4}{2\Lambda(\sigma_{\text{F}}^2 + 4\Lambda^2\Sigma_V^2)},\\ &= \frac{1}{2}\sigma_{\text{F}}\Sigma_V.\end{aligned}$$

Therefore, the conditional expectation of Alice's gain increases with the amount of private information and with the overall liquidity of the market, which is (unwittingly) provided by the noise traders. For typical values of the predictor (i.e. for values of $(p_{\text{F}} - p_0)$ of the order of σ_{F}), it follows that $\widehat{Q} \sim \Sigma_V$, so Alice contributes a substantial fraction of the total traded volume. This is not very realistic in practice: as a common-sense precautionary measure, informed traders tend to limit their trading to a small fraction of the total volume. Out-sized trades risk destabilising the market, resulting in a much larger impact cost (see discussion in Section 10.5).

Finally, the pricing error at $t = 1$ can be measured as:

$$\mathbb{V}[\widehat{p} - p_\mathrm{F}] = \frac{1}{2}\mathbb{V}[p_\mathrm{F} - p_0] = \frac{1}{2}\sigma_\mathrm{F}^2. \tag{15.8}$$

Therefore, Bob is only able to reduce by one-half the variance of the uncertainty of the fundamental price known to Alice.

The Kyle model provides a clear picture of the origin of price impact. In the model, market-makers fear that someone in the market is informed, and therefore react by increasing the price when they observe a surplus of buyers and decreasing the price when they observe a surplus of sellers. The model also illustrates that even though the model permits Alice to enter an infinitely large position, her private information only leads to bounded profits, because of impact costs.

Note that when moving slightly away from its core assumptions, the Kyle model becomes a self-fulfilling mechanism. Suppose that market-makers overestimate σ_F (i.e. they overestimate the quality of information available to insiders). In reality, because there is no "terminal time" when a "true price" p_F is revealed, such market-makers will over-react to the order flows when setting the value of $\widehat{p}$. In the efficient-market picture, these pricing errors should self-correct through arbitrage from the informed trader (i.e. with a signal to trade in the opposite direction at the next time steps). This would lead to excess high-frequency volatility. However, signature plots of empirical data are rather flat (see Section 2.1.4), which instead suggests that the whole market shifts its expectations around the new traded price, much as assumed in the Santa Fe model in Chapter 8 (for which $\sigma_\mathrm{F} = 0$, which corresponds to the situation of no information but non-zero impact).

15.4 Some Extensions

There are several simple ways to extend the Kyle model:

(i) One possible extension, due to Kyle himself, is to consider a multi-step set-up in which the terminal price p_F known to Alice only reveals itself at time T. In the continuous-time limit, in which there are infinitely many steps between the initial time and the terminal time T, Alice's optimal trade at any time $t < T$ is still linear in the mispricing:

$$\varepsilon(t)\,\mathrm{d}\widehat{Q}(t) = \frac{\Sigma_V T}{\sigma_\mathrm{F}(T-t)}(p_\mathrm{F} - p(t))\,\mathrm{d}t, \tag{15.9}$$

while the impact parameter is independent of time and given by $\Lambda = \sigma_\mathrm{F}/\Sigma_V$. The resulting price volatility is also constant in time.[1] Quite remarkably,

[1] The fact that $\Lambda(t)$ is a constant in a continuous auction equilibrium implies that trading prices have constant volatility over time and therefore that information is gradually incorporated into prices at a constant rate.

 https://doi.org/10.1017/9781316659335.023

due to the market-maker's clearing rule, even in the presence of Alice's systematic trading in the direction of the true price p_{F}, the price $p(t)$ remains a martingale for Bob. A similar property will hold in the Glosten–Milgrom model (see Section 16.1.5). Contrarily to empirical observations, however, the multi-time-step Kyle model's order-sign series is not autocorrelated at all!

In this model, Alice's aggressiveness increases as $t \to T$. In order to maximise her profit, Alice's impact on the price ensures that $p(t) \to p_{\mathrm{F}}$ as $t \to T$. This is intuitively obvious if Alice's information is certainly true, because any mispricing would lead to unexploited profit opportunities.

Strangely, however, this convergence is actually driven by Alice's trading strategy, Equation (15.9). In other words, if Alice firmly believed that the price at time T should be equal to some arbitrary value p_{Alice} and Bob again acted as if Alice was truly informed, then the price would indeed converge to p_{Alice} at time T.

(ii) Another possible extension is to assume that Alice is risk-averse, so adds a risk-penalty term of the form $-\zeta Q^2$ to her objective function, such that Equation (15.3) instead becomes

$$\mathcal{G} = \varepsilon Q \times (p_{\mathrm{F}} - p_0 - \varepsilon \Lambda Q) - \zeta Q^2.$$

To leading order in ζ, the inclusion of this term decreases the value of Kyle's lambda to $\Lambda - \zeta^2/2\Lambda$. Since Alice's new risk-aversion term constrains her from taking big positions, the adverse selection risk is reduced and Bob can offer more liquidity to the market.

(iii) A third simple extension is to remove the hypothesis that V_{noise} and $p_{\mathrm{F}} - p_0$ are Gaussian random variables. In this case, one can still solve Kyle's model in the small-volume limit $Q \ll \Sigma_V$. Provided that:

- the distribution of the random component of the order flow has a quadratic maximum around zero, i.e. it behaves as

$$\mathbb{P}(V_{\mathrm{noise}} \to 0) = P_0 - P_0'' V_{\mathrm{noise}}^2/2 + \dots;$$

- the distribution of $p_{\mathrm{F}} - p_t$ has a finite variance σ_{F}^2,

then the main conclusions of the Kyle model still hold, in particular that price impact is linear when $Q \ll \Sigma_V$, with a coefficient proportional to σ_{F}.

Besides these three, many other extensions and generalisations of the Kyle model have been considered, for example the role of inventory risk constraints for the market-maker.[2]

15.5 Conclusion

The Kyle model elegantly elicits some deep truths about how markets function, but also fails to capture some important empirical properties of real markets. The most interesting outcomes of the model are:

- Trades impact prices. In the model, the mechanism that creates price impact is the market-maker's attempts to guess the amount of information contained in the order-flow, and to adjust the price up or down accordingly.

[2] See: Cetin, U., & Danilova, A. (2016). Markovian Nash equilibrium in financial markets with asymmetric information and related forward-backward systems. *The Annals of Applied Probability*, 26(4), 1996–2029.

- Impact is linear (i.e. price changes are proportional to order-flow imbalance) and permanent (i.e. there is no decay of impact). In the context of an LOB, linear impact is a generic consequence of having a finite density of buy/sell orders in the vicinity of the price, and permanent impact is a consequence of liquidity immediately refilling the gap left behind the incoming market order.[3] We will return to this discussion in the context of Walrasian auctions in Chapter 18.
- The impact of trading is inversely proportional to the total volume traded by noise traders. In other words, the existence of uninformed traders is essential for the market to function. In the next chapter, we will see that the Glosten–Milgrom model reaches a similar conclusion: a dearth of uninformed trading can lead to market breakdown.
- Because of impact, the informed trader must limit her trading volume to optimise her gains. Therefore, the amount of profit that can be made using private (insider) information is limited.

When confronted with empirical data, the Kyle model suffers from important drawbacks. The most obvious one is that the sign of order flow is found to be completely uncorrelated. This is a consequence of impact being permanent in the Kyle model. Any correlation in signs would lead to predictable returns, as emphasised in Sections 10.3 and 13.1. Therefore, in order to capture realistic market dynamics, the Kyle framework must be extended to accommodate sign correlations – a topic that we discussed in Chapter 13 and to which we will return in Section 16.2.1 below.

The fact that impact in the Kyle model is linear and permanent is at odds with the square-root impact law of Chapter 12. It also leads to the conclusion that returns and aggregate volume imbalance are related through the very same constant Λ, independently of the time scale T over which they are computed. This is in strong contrast with the empirical data shown in Section 11.4, which suggests that $\Lambda(T)$ in fact decays as $\sim T^{-0.25}$. Clearly, additional features must be included in Kyle's model to make it a realistic model of price impact.

Take-Home Messages

(i) The Kyle model is a simple model of price impact with three classes of agents. The informed trader has private information about the future price, and chooses a trade volume to optimise her profit. The noise traders submit a random trade volume. The market-maker acts as a counterpart for the sum of the trading volumes submitted by the informed trader and the noise traders, and chooses his clearing price

[3] See also Obizhaeva, A., & Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16, 1–32.

to equal his expectation of the fundamental price, given the volumes he observes – such that his expected profit is zero.

(ii) The informed trader anticipates the market-maker's clearing price and therefore optimises her volume to maximise her expected profit.

(iii) In the model, volumes impact the price because of their expected informational content. This exposes the market-maker to adverse selection. By adjusting the price (negatively) to order-flow imbalance, the market-maker ensures that on average, impact exactly compensates for this adverse selection.

(iv) When all distributions are Gaussian, the impact scales linearly with the informed trader's volume. The proportionality coefficient is often called Kyle's lambda.

(v) The value of Kyle's lambda measures market (il-)liquidity. The larger the coefficient, the more a given volume impacts the price and the more expensive trading is.

(vi) The larger the number of noise traders, the more liquid the market is. In this sense, a market needs uninformed participants to function smoothly.

15.6 Further Reading

The Kyle Model and Some Generalisations

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 53, 1315–1335.

Back, K. (1992). Insider trading in continuous time. *Review of Financial Studies*, 5(3), 387–409.

Caballe, J., & Krishnan, M. (1994). Imperfect competition in a multi-security market with risk neutrality. *Econometrica*, 695–704.

Nöldeke, G., & Tröger, T. (2001). Existence of linear equilibria in the Kyle model with multiple informed traders. *Economics Letters*, 72(2), 159–164.

Corcuera, J. M., Farkas, G., Di Nunno, G., & Oksendal, B. (2010). Kyle-Back's model with Lévy noise. Preprint series. Pure mathematics. http://urn. nb. no/URN: NBN: no-8076.

Cetin, U., & Danilova, A. (2016). Markovian Nash equilibrium in financial markets with asymmetric information and related forward-backward systems. *The Annals of Applied Probability*, 26(4), 1996–2029.