Trading with Market Impact

H. Mete SONER,
Department of Mathematics ETH Zürich
Swiss Finance Institute

Mathematical Finance beyond classical models
Zurich, September 16, 2015

I have benefited the collaboration of many people including : Albert Altarovici, Peter Bank, Umut Çetin, Yan Dolinsky, Selim Gökay, Ludovic Moreau, Dylan Possamaï, Max Reppen, Alexandre Roch, Moritz Voss and

Nizar Touzi

Johannes Muhle-Karbe

Consider a financial market in which our trades impact the current value of the stock. We would like to

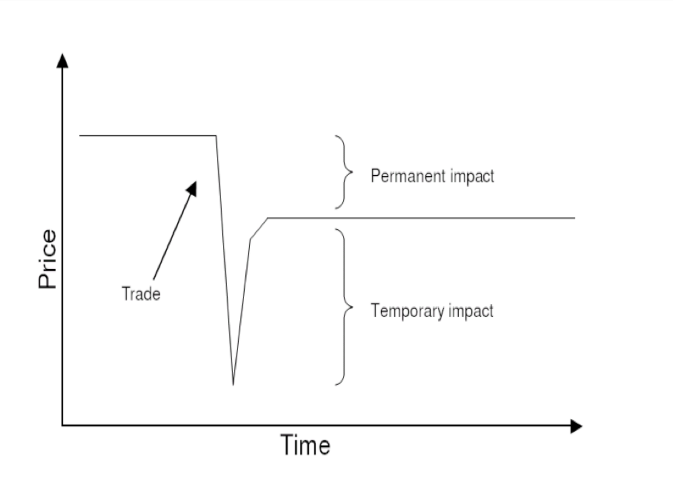
- model the market impact,
- model the dynamics of this impact,
- study its impact on investment decisions.

- These models involve a large trader placing substantially big orders. This results in a transformation of, in addition to price, the volatility of the security. In particular, the volatility of the transformed process can become time and size dependent.
- In turn, they cause an permanent impact on the price.
- I will not consider these models in this talk and refer you to Jarrow, Frey & Stremme, Frey, Huberman & Stanzl, Cvitanic & Ma, Cvitanic & Cuoco, Platen & Schweizer, Roch.

- Çetin-Jarrow-Protter model of liquidity is the representative for this type of models. In this setting the authors postulate the existence of a supply curve for the price process of the asset.
- The supply curve gives you the price per share once you specify the time and size of the trade.
- All investors are price takers to the supply curve and have no lasting impact on the evolution of the underlying.

Even in temporary impact models, it is realistic to assume that the trade impacts the price. The main feature of a temporary impact model is that the change in price should disappear if there is no trade in the near future. In summary,

- When a trade of size ΔZ_t is made, price changes by a certain function of ΔZ_t;
- Then, a certain fraction of this change remains (permanent impact);
- The remaining price change (temporary impact) decays in time (relaxation).


Price
Trade
Permanent impact
Temporary impact
Time

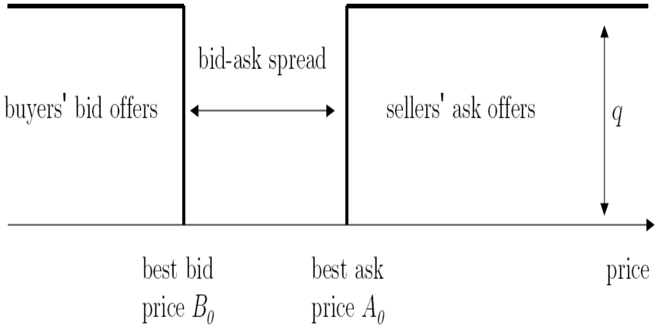


Fig. 1. The limit order book model before the large investor is active

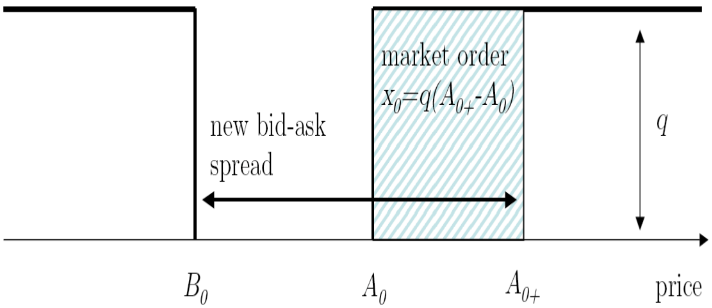


Fig. 2. Impact of a market buy order of x_0 shares

Any model with friction is somehow related to this problem of price impact. Indeed, transaction costs can be taught as a particular price impact. Any positive amount of trade pushes the price to the ask-price and any sale to the bid.

In the context of hedging, Leland formally argued that transaction cost modifies the volatility. Later, Fukasawa and Rosenbaum & Tankov revisited this approach. To understand this modification, jointly with Barles we used asymptotics to obtain a modified Black & Scholes equation.

We may ask many relevant questions in such a model. First interesting one is how to optimally execute a large order. In practice this is quite important.

In general these orders are split into smaller pieces and executed over time. Then, one has to balance between the need and the risk of splitting into too many pieces and waiting a longer time and placing large orders and paying higher cost of liquidity.

In this talk, I will not treat this problem *directly*. See Bertsimas & Lo, Almgren, Almgren & Chriss, Schied & Schöneborn, Gatheral, Obizhaeva & Wang, Alfonsi, Fruth & Schied and Alfonsi & Schied for more information.

Price Impact

Related

Cetin-Jarrow-Protter Model

Supply Curve

CJP Dynamics

Super-replication for CJP

Penalizing the speed

Almgren-Chriss

Asymptotics

Hedging

They postulate an exogenous supply curve

$$\boxed{\mathcal{S}(t, S_t, \nu), \qquad S_t = \mathcal{S}(t, S_t, 0)}$$

which gives the price per share for a transaction of size ν ($\nu > 0$ is a buy and $\nu < 0$ a sell). An example of the supply curve is the generalized Black-Scholes economy with liquidity parameter Λ :

$$\mathcal{S}(t, S_t, \nu) = S_t \exp\left(\Lambda \nu\right), \quad dS_t = S_t \left[\mu dt + \sigma dW_t\right].$$

We may simply take

$$\boxed{\mathcal{S}(t, S_t, \nu) = S_t + \Lambda\nu.}$$

This corresponds to a constant density LOB. And $1/\Lambda$ is the constant density. For a transaction of size ν we pay

$$\boxed{\nu S_t + \Lambda\, \nu^2 = \nu\, [\mathcal{S}(t, S_t, \nu) - S_t]\,.}$$

Now imagine of splitting this order into two and execute them in tandem. Then we would pay

$$2\ \left[(\nu/2)S_t + \Lambda((\nu/2))^2\right] = \nu S_t + \frac{1}{2}\Lambda\, \nu^2.$$

Hence in this model,we would like to make small but many transactions. Hence the portfolio process should be continuous with no liquidity cost. But :

- In reality, each transaction takes some time to execute - finite speed of the portfolio. So, we carry the risk of stock price movements.
- Secondly, after the first transaction stock price moves and we will not be able to get the same price as before - i.e. there is resilience. But in the ideal model of Cetin, Protter and Jarrow, we do get the same price !

- Jointly with Çetin & Touzi we propose to restrict the portfolio to be a semimartingale.
- Jointly with Gökay and later with Dolinsky, we consider the limit of discrete time markets.
- Jointly with Roch and later Vukelja we introduce exponential relaxation into the model.
- In a slightly different and more phenomenological model, Almgren & Chriss penalises the square of the speed of change of the portfolio. Later, jointly with Moreau & Muhle-Karbe, we analysed this model asymptotically.
- Jointly, with Bank & Voss, we propose an ad-hoc model to construct hedges.

ETH zürich

swiss:finance:institute

As usual one risky and one riskless asset and with zero interest rate.

$(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space.

As usual one risky and one riskless asset and with zero interest rate.

$(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space. Let X_t be the units of money-market account, the semimartingale Z_t be the number of shares. We *define* that Z is a *self-financing* strategy if

$$
\begin{aligned}
Y_t \; &:= X_t + Z_t S_t = X_0 + \int_0^t Z_t dS_t \\
&\quad - \sum_{0 \le u \le t} \Delta Z_u \left[\mathcal{S}(u, S_u, \Delta Z_u) - S_u \right] - \int_0^t \frac{\partial \mathcal{S}}{\partial \nu}(u, S_u, 0) d[Z, Z]_t^c.
\end{aligned}
$$

This derivation is done first for elementary processes and then passing to limit.

$$dS_t = S_t\left[\mu dt + \sigma dW_t\right]$$

$$Y_t = +\int_0^t Z_u dS_u - \sum_{n=0}^{N-1} z_n \left[\mathcal{S}\left(\tau_n, z^n\right) - \mathcal{S}\left(\tau_n, 0\right)\right] \mathbf{1}_{\{t<\tau_{n+1}\}}$$

$$- \int_0^t \frac{\partial \mathcal{S}}{\partial \nu}(u, S_u, 0)\, \Gamma_u^2 \sigma^2 S_u^2 du.$$

$$Z_r = \sum_{n=0}^{N-1} z_n \mathbf{1}_{\{t<\tau_{n+1}\}} + \int_0^t \alpha_u du + \int_0^t \Gamma_u dS_u.$$

For a continuous Z (i.e., no “large trades”), we still have a liquidity cost in terms of the Gamma of the portfolio.

On this supply curve there is no unique value of the portfolio, for instance one can consider the liquidation value $X_t + Z_t\mathcal{S}(t, -Z_t)$ or the book value of the portfolio $Y_t = X_t + Z_t S_t$.
We use the book value :

$$Y_t = \int_0^t Z_u dS_u - L_t$$

$$L_t = \sum_{0\leq u\leq t} \Delta Z_u[\mathcal{S}(u, S_u, \Delta Z_u) - S_u] + \int_0^t \frac{\partial \mathcal{S}}{\partial \nu}(u, S_u, 0)\Gamma_u^2\sigma^2 S_u^2 du.$$

Since $\mathcal{S}$ is monotone in ν, $L_T \geq 0$. Notice that that if Z is of finite variation and continuous, then $L_t = 0$.

Consider a process Z and its stochastic integral,

$$C := \int_0^T Z_u dS_u.$$

Leventhal & Skorokhod and Bank & Baum proved that one can construct a sequence self-financing trading strategies Z^n's that are absolutely continuous (hence $L^n_t = 0$) and

$$X^n_T := \int_0^T Z^n_u dS_u - \underbrace{L^n_T}_{=0} \to \int_0^T Z_u dS_u = C.$$

uniformly.

The construction is through a Borel-Cantello argument and the BV norm of the approximating portfolio Z^n gets arbitrarily large.

Simply, we restrict the portfolio and its gamma to be a semimartingales with some bounds on the characteristics.

Let $\mathcal{A}_{t,s}$ be the set of all admissible portfolios.

For a given a European contingent claim with payoff g, the super-replication cost is defined by

$$\boxed{V(t,s) = \inf\left\{ y : Y_T^{t,y,Z} \geq g(S_T^{t,s}) \text{ } a.s. \text{ for some } Z \in \mathcal{A}_{t,s} \right\}}$$

Together with Cetin and Touzi, we showed that the super-replicating cost satisfies

$$0 = -V_t + \sup_{\beta \geq 0} \left(-\frac{1}{2} s^2 \sigma^2 (V_{ss} + \beta) - \Lambda s^2 \sigma^2 (V_{ss} + \beta)^2 \right),$$

together with terminal cost $V(T, s) = g(s)$. We rewrite as

$$-V_t - s^2 \sigma H(V_{ss}) = 0,$$

where the function H is given by

$$H(\gamma) = \begin{cases} \frac{1}{2}\gamma + \Lambda\gamma^2 & \gamma \geq -\frac{1}{4\Lambda} \\ -\frac{1}{16\Lambda} & \gamma \leq -\frac{1}{4\Lambda} \end{cases}.$$

For a convex pay-off g, the solution also remains convex and the equation simplifies to,

$$\begin{aligned} V_t &= -\frac{1}{2}s^2\sigma^2 V_{ss} - \Lambda s^2\sigma^2 (V_{ss})^2, \\ &= -\frac{1}{2}s^2\hat{\sigma}^2(t,s) V_{ss}, \end{aligned}$$

where

$$\boxed{\hat{\sigma}^2(t,s) = \sigma^2\left[1 + 2\Lambda V_{ss}(t,s)\right].}$$

Hence the effect of liquidity is to increase the *effective volatility* as in Leland and Fukasawa. Same is true for non-convex pay-off’s as well.

- By an application of maximum principle we have

$$V(t,s) \geq V_{BS}(t,s)$$

  and they are equal only when g is an affine function.
- This implies that there exists a strict liquidity premium, a difference between the superreplicating cost and the Black-Scholes value of the claim.
- The reason why there are contradicting results between CJP and the above is the trading strategy constraints.

- Together with Gökay and later with Dolinsky for general pay-offs, we consider the discrete-time super-replication problem, where we do not impose any conditions on the portfolio processes.
- We analyze the asymptotic limit of the binomial model numerically and theoretically. We find out as time steps gets smaller, we recover the same PDE as in the portfolio constrained case.
- This justifies the necessity of the constraints on the portfolio strategies in the continuous time paper.

Convex dual representation is crucial in Dolinsky & Soner.

- CJP model has liquidity premium.
- This premium is due to the resilience and can be realised by restrictions on the portfolio.
- Supe-replication is cost is different than the frictionless model.
- However, without introducing the resilience explicitly, the liquidity premium is weak and is does not impact the utility maximization problems.

This is a phenomenological model by Almgren & Chriss (also important contributions by Rogers & Sign, Garleanu & Pedersen), considers an impact functional of the form

$$\mathcal{S}(t, S_t, Z'_t) = S_t + \Lambda Z'_t.$$

Then, the dynamics are given by

$$Y_t = \int_0^t Z_u dS_u - L_t$$

$$L_t = \Lambda \int_0^t \left(Z'_u\right)^2 du.$$

In these models, it is not possible to avoid the liquidity premium.

Consider a utility maximization problem

$$\boxed{\sup_{Z} \mathbb{E}\left[\, U\left(\mathcal{R}_T^Z\right)\right],}$$

where $\mathcal{R}_T$ is the risk adjusted liquidation cost of Schöneborn and is given by,

$$\mathcal{R}_T^Z := Y_T^Z - C\Lambda^2(Z_T - Z_T^*)^2,$$

where C is a constant derived from the model and Z^* is optimal portfolio for the frictionless (i.e., $\Lambda = 0$) market.

There are two difficulties :

- Due to the price impact, we could only use portfolios that are differentiable in time. If the target portfolio Z^* is rough, the optimisation problem gives us a way to approximate this target portfolio.
- In addition to continuous targeting error, we have both initial and final liquidation costs.
- Initially, we might far from the optimal location and need to move there efficiently.
- Also, closer to maturity one must consider the final portfolio position.

The actual problem is not quite tractable and together with Moreau & Muhle-Karbe we considered the asymptotics as Λ gets smaller.

We have asymptotic results for the value function and also for optimal portfolio.

The rigorous proof uses recent machinery from viscosity solutions which I do not report here. Only I outline the asymptotic structure of the hedge.

Let $z^* = z^{*,\Lambda}$ be the optimal portfolio for the utility maximization problem with small but non-zero impact $\Lambda > 0$.

Asymptotically,

$$\boxed{\frac{d}{dt} z^*_t = c\,\Lambda^{-1/2}\,(z^*_t - Z^*_t), \quad \text{where} c = \frac{\sigma}{\sqrt{2R_t}},}$$

and R_t is the frictionless investor's indirect risk-tolerance process, i.e., the risk tolerance of the frictionless value.

As Λ gets smaller, z^* moves very quickly towards the frictionless optimizer Z^*.

- A related model is the friction due to transaction costs. Asymptotics analysis has been successfully used in that context by Shreve and collaborators, by myself with Altarovici, Reppen, Muhle-Karbe, Touzi.
- Relatedly, in a series of papers Kallsen & Muhle-Karbe studied directly the asymptotics of the optimal portfolio.
- Kallsen & Muhle-Karbe formulae although different in their fine details, have many common features. In particular, the risk tolerance function plays a central role. But the scaling in the small parameter might be different.

Joint with Bank & Voss, we consider the following tracking problem for a given portfolio process Z^*_t,

$$\text{minimize} \quad J(u) := J(u; x, Z^*),$$

where

$$J(u; x, Z^*) := \frac{1}{2} \int_0^T \left[\left(X_t^{u,x} - Z_t^* \right)^2 + \Lambda u_t^2 \right] dt,$$
$$X_t^{u,x} := X + \int_0^t u_s ds.$$

The above model is motivated by recent papers of Bank & Voss and also Kallsen & Muhle-Karbe. It was also considered in Rogers & Sign but solved only approximately.

The optimizer X^* has a very similar structure to the asymptotic formula already discussed in the impact model. Indeed, it solves

$$\boxed{\frac{d}{dt}X_t^* = c\,\Lambda^{-1/2}\,(X_t^* - (\mathcal{L}Z^*)_t),}$$

where $\mathcal{L}Z^*$ is a linear map of Z^* depending on the parameter Λ. Roughly, it is the adapted projection of the forward convolution of Z^*.

So, instead of targeting directly the target portfolio Z_t^* at time t, we target an estimate of the possible future values of the target. This was also obtained by Garleanu & Pedersen.

Garleanu & Pedersen quote Wayne Gretzky, *“A great hockey player skates to where the puck is going to be, not where it is.”*

- There are a rich class of models for illiquid markets with price impact.
- Another use of this approach is to assume that the target portfolio is given but not implementable. This would give us away to provide implementable approximations.
- Asymptotics makes things tractable.

THANK YOU FOR YOUR ATTENTION.