Imperial College
London

# Market Microstructure
# Problem Set 2

Wednesday 5 February 2020
Johannes Muhle-Karbe, j.muhle-karbe@imperial.ac.uk
Claudio Bellani, c.bellani17@imperial.ac.uk

**Exercise 1** ("Imperfectly informed investors in the Glosten-Milgrom model")**.** We would like to modify Glosten-Milgrom model by relaxing the dichotomy informed/noise traders. More precisely, we would like to say that the informed traders are in fact traders who utilise a "signal" to know the true value $\nu$ of the traded asset; this signal can be accurate with some probability $1/2 < \rho \leq 1$, but it can also be inaccurate (i.e. misleading) with probability $1-\rho$. If the signal equals $\nu^H$, traders who rely on the signal buy the security; if the signal equals $\nu^L$, they sell. Hence, if the signal is accurate, their interventions in the market are essentially like those of perfectly informed traders in the classical Glosten-Milgrom model presented in lecture slides (they make gain at the expense of market makers); if instead the signal happens to be inaccurate, they take the unfavourable side of the trade, making a loss. We call this type of traders *imperfectly informed*.

This exercise requires you to compute the optimal quotes posted by perfectly competitive risk-neutral market makers who provide liquidity to either imperfectly informed traders or noise traders.

1. Reasoning as in the classical Glosten-Milgrom model, compute a market maker's expected profits when she receives a buy order and when she receives a sell order. [Hint: The profits depend on the market maker's counterpart; if her counterpart is an imperfectly informed trader, she will be facing a better-informed counterpart with probability $\rho$.]
2. Compute the ask and the bid prices that a risk neutral market maker quotes.
3. Express the bid-ask spread as a function of the signal accuracy $\rho$. Compare this with the classical model with perfectly informed investors.
4. Compute the expected gain made by imperfectly informed traders who rely on the signal when trading. Is this expected gain positive?

**Exercise 2** ("Implementation of Glosten-Milgrom model")**.** This exercise requires you to implement the version of Glosten-Milgrom model presented in the lecture slides. The dataset for the implementation is the same used in Problem Set 1. It consists of the two csv files

```
MSFT_2012-06-21_34200000-57600000_message_10.csv
MSFT_2012-06-21_34200000-57600000_orderbook_10.csv
```

that you can download from the the Blackboard folder `Assignment_1`. From the lecture slides, recall that the theoretical ask prices computed in the Glosten-Milgrom model are given by the formulae

$$\begin{aligned} a_{t+1} =& \mu_t + \frac{2\pi\theta_t(1-\theta_t)}{1+\pi(2\theta_t-1)}(\nu^H - \nu^L), \\ a_1 =& \mu_0 + \frac{\pi}{2}(\nu^H - \nu^L), \end{aligned} \tag{1}$$

with symmetric counterparts for the bid prices $b_t$. Recall also the interpretation of the variables appearing on the right-hand side of equation (1): $\nu^L < \nu^H$ are the two possible values of the "true" price $\nu$ of the traded asset; $\theta_{t+1} = \mathbb{P}(\nu = \nu^H | \mathfrak{F}_t, d_{t+1})$ is the probability that the dealers assign to the event $\{\nu = \nu^H\}$ at time $t$ conditioned on the direction of the $t+1$-th trade; $\mu_t = \mathbb{E}[\nu | \mathfrak{F}_{t-1}, d_t] = \theta_t \nu^H + (1-\theta_t)\nu^L$ is the expected value of $\nu$ at time $t-1$ conditioned on the $t$-th trade; $\pi$ is the the probability that an order is placed by some informed trader. The update rule for $\theta_{t+1}$ is

$$\begin{aligned} \theta_{t+1} =& \frac{(1+\pi d_{t+1})\theta_t}{1+(2\theta_t-1)\pi d_{t+1}}, \\ \theta_0 =& 1/2, \end{aligned}$$

where $\pi$ is the the probability that an order is placed by some informed trader.

Let $\Delta\nu$ be the difference $\nu^H - \nu^L$. The exercise requires you to estimate $\pi$ and $\Delta\nu$ from the orderbook data of MSFT. Proceed as follows:

1. Consider the rows of `MSFT_2012-06-21_34200000-57600000_message_10.csv` that correspond to events of type 4 (i.e. executions). If $n$ is the index of one of these rows, extract the $n-1$-th row from `MSFT_2012-06-21_34200000-57600000_orderbook_10.csv`, for all $n$. In this way you extracted from the message file the subset corresponding to trades that happened on 21 June 2012, and from the orderbook file the states of the LOB immediately before these trades. Store these two subsets of the original datasets under the names `trades.csv` and `lob_trades.csv`.
2. Let $d_t$ be the directions of the transactions reported in `trades.csv`, and let $a_t^{\text{data}}$ and $b_t^{\text{data}}$ be the ask and the bid prices reported in `lob_trades.csv`.
3. Estimate $\pi$ and $\Delta\nu$ by minimising the residual sum of squares
$$(\pi, \Delta\nu) \mapsto \sum_t (a_t^{\text{data}} - a_t)^2 + \sum_t (b_t^{\text{data}} - b_t)^2,$$
$$0 < \pi < 1, \quad \Delta\nu > 0,$$
where $a_t$ and $b_t$ are computed from equation (1) by using the sequence $(d_t)$ of trade directions reported in `trades.csv`.
4. Once the estimated values $\pi$ and $\Delta\nu$ have been obtained, "replay" the trajectories of $t \mapsto a_t$ and $t \mapsto b_t$ and compare these with the real ones $t \mapsto a_t^{\text{data}}$ and $t \mapsto b_t^{\text{data}}$ by plotting them in the same graph. How good (or bad) the fit is?
5. Consider the theoretical spread $S_t = a_t - b_t$ and the real-world spread $S_t^{\text{data}} = a_t^{\text{data}} - b_t^{\text{data}}$. How do the two compare to each other? In the theoretical model, do you observe the fact that over time the information about the true value of $\nu$ is progressively revealed? What would you add to the model to prevent this from happening and (intuitively) why?

# Solutions

**Solution to Exercise 1.** A market maker's expected profit when she receives a buy order at time $t+1$ is
$$\Big(a_{t+1} - \mu_t\Big)\frac{1-\pi}{2} + \Big(a_{t+1} - \nu^H\Big)\pi\theta_t\rho + \Big(a_{t+1} - \nu^L\Big)\pi(1-\theta_t)(1-\rho).$$
A market maker's expected profit when she receives a sell order at time $t+1$ is
$$\Big(\mu_t - b_{t+1}\Big)\frac{1-\pi}{2} + \Big(\nu^L - b_{t+1}\Big)\pi(1-\theta_t)\rho + \Big(\nu^H - b_{t+1}\Big)\pi(\theta_t)(1-\rho).$$
This answers the first point. Equating these two formulas to zero and solving for $a_{t+1}$ and for $b_{t+1}$ gives the risk neutral quotes, answering the second point. With these quotes, simply compute the difference $a_{t+1} - b_{t+1}$ to answer the third point. Finally, when imperfectly informed traders buy, they expect a profit of
$$\Big(\nu^H - a_{t+1}\Big)\rho + \Big(\nu^L - a_{t+1}\Big)(1-\rho),$$
and when they sell they expect a profit of
$$\Big(b_{t+1} - \nu^L\Big)\rho + \Big(b_{t+1} - \nu^H\Big)(1-\rho).$$
Plug the risk neutral quotes $a_{t+1}$ and $b_{t+1}$ in these two expectations to answer the fourth point.