Stock price dynamics and the bid-ask spread from the Market Maker’s perspective

Diego Gomez
Mehdi Talbi

2020-2021
Contents

1 Introduction .................................................. 2

2 Related work ................................................... 3
   2.1 Financial terminology .................................... 3
   2.2 Brownian Motion ......................................... 4
   2.3 Re-scaled Random Walk ................................... 6
   2.4 The Cox-Ross-Rubinstein model ......................... 7

3 Problem ......................................................... 8
   3.1 Motivation ................................................. 8
   3.2 Formalization .............................................. 8
   3.3 Intuition of the problem .................................. 10

4 Theoretical work .............................................. 12
   4.1 1-period problem ......................................... 12
   4.2 2-period problem ......................................... 13
   4.3 Derivation of the recursive identity ..................... 16

5 Numerical Solution .......................................... 18
   5.1 Defining the grid ........................................ 20
   5.2 Results .................................................. 21

6 Future Work .................................................. 24

References ....................................................... 25
1 Introduction

The traditional trading venues are being challenged by the emergence of alternative ones. This mainly due to the fragmentation of financial markets, which incites competition in the exchange of financial assets \[3\]. Therefore pricing assets in an optimal way is of ever-growing importance for these financial institutions. That is because while such agents seek to maximize their returns by increasing the gap between the price of the asset and the bid and ask prices, other competitors can now easily outmatch these gaps, consequently attracting potential clients.

We are inspired by \[2\] where the authors decided to tackle the problem of finding optimal make-take fees using a continuous model. In the following paper, we will be tackling the problem of finding the optimal bid and ask prices for a given financial asset. More specifically in this Bachelor Thesis, we are interested in the market maker’s perspective on this problem. That is if we are the ones deciding the prices of the stock at each given time, what are the prices for which the stock yields the biggest return and attracts the most clients. In the following, we will propose a solution to the discrete version of the problem, as it better suits reality, and it is simpler to solve.

The paper is organized as follows. We start by showcasing the notions that are necessary for the proper understanding of this paper and the ones that helped us tackle the problem. We then proceed to motivate and formalize the problem we seek to solve. Then we present the theoretical work done to solve and further our understanding of the problem. Next, we present our approach we took to solving it as well as the results obtained. Finally, we finish this paper with a proposal of what could be the next step to further this research.
2 Related work

In order to tackle this problem we need to start by properly defining some financial terms, and stating the main mathematical objects and theorems on which this paper is based. Please note that the reader is expected to be familiar with probability theory, mainly the usual distributions, as well as conditional expectations.

2.1 Financial terminology

First we state the financial notions required for the problem.

Definition 2.1.1

1. A zero-intelligence model, is a model where we assume that agents act randomly as opposed to agents acting strategically to optimize their portfolio.\[1\]

2. The ask-price is the price at which an agent of the market can buy a given stock.

3. The bid-price is the price at which an agent of the market can sell a given stock.

4. The bid-ask spread is the difference between the ask-price and the bid-price.

5. The market maker is the agent setting the bid and ask prices of an asset, usually hoping to make profit of the bid-ask spread.

As said before, we will consider the point of view of the market maker in the following problem.
2.2 Brownian Motion

This phenomenon was named after Robert Brown the Scottish botanist who first described it while observing the movement of pollen suspended in water. Since its first description, this pattern of motion has proved to be useful to model several irregular movements in many different fields, finance included.

It was in his Thesis in 1900 that Louis Bachelier first proposed to use the Brownian Motion to model Stock Prices. Nevertheless, his work was not recognized until recent history. It was Fisher Black, Myron Scholes, and Robert Merton that realized the real success of the Brownian Motion, and consequently received the Nobel Prize in Economics in 1997 founding the modern theory of financial mathematics. Finally, this phenomenon is at times also referred to as a Wiener process, because Norbert Wiener first constructed rigorously this motion.

Let us now state the definition of this phenomenon.

**Definition 2.1.1** A stochastic process \( B := (B_t)_{t \geq 0} \) is called a Brownian motion if it satisfies the following

1. \( B_0 = 0 \)

2. With probability 1, the paths of \( B \) are continuous that is,

\[
P[\{\omega, t \to B_t(\omega) \text{ is continuous }\}] = 1
\]

3. The time increments of \( B \) are independent, that is for any \( 0 \leq t_i < t_{i+1} \leq t_j < t_{j+1}, \)

\[
B_{t_{i+1}} - B_{t_i} \text{ is independent from } B_{t_{j+1}} - B_{t_j}
\]

4. For any \( 0 \leq s \leq t \), the random variable \( B_t - B_s \sim \mathcal{N}(0, t - s) \)
For example, we can plot a simulated Brownian Motion from time 0 to time 1.

![Brownian Motion](image)

Based on this, one can define a very interesting stochastic process, that is partially deterministic and partially random. We call this an arithmetic Brownian Motion.

**Definition 2.1.2** The stochastic process \((X_t)_{t \geq 0}\) is called an Arithmetic Brownian motion if it satisfies the following Stochastic Differential Equation (SDE)

\[
dX_t = mdt + \sigma dB_t
\]

with \(B\) a Brownian Motion. We call \(m\) the drift, and \(\sigma\) the volatility.

It can be more intuitive to see the solution of this SDE

\[
X_t = x_0 + mt + \sigma B_t \text{ with, } x_0 \in \mathbb{R}
\]

Now we see that an Arithmetic Brownian Motion is nothing but a Brownian Motion summed to an affine function.
2.3 Re-scaled Random Walk

The random walk is a stochastic process with applications ranging from physics to financial mathematics as we are about to see. One of the powerful properties of this process is its relation with the previously defined Brownian Motion.

**Definition 2.3.1** The random walk on $\mathbb{Z}$ is defined as

$$X_0 = 0, X_n = \sum_{m=1}^{n} Y_m, n \in \mathbb{N}^*$$

with $(Y_m)_{m \in \mathbb{N}^*}$ a sequence of i.i.d random variables such that, $\mathbb{E}[Y_m] = 0$ and $[Y_m] = 1$.

Indeed, thanks to the plot we have of the Brownian Motion, the fact that there exists a very strong relationship between these processes becomes less surprising.

**Proposition 2.3.1** Defining $(X_n), (Y_n)$ as in the previous definition.

$$B_n(t) = \frac{X_{\lfloor nt \rfloor}}{\sqrt{n}}$$

converges in law to a random variable $\tilde{B}_t$ with the stochastic process $\tilde{B}$ satisfying properties 1, 3, 4 of the Brownian Motion definition. Moreover since the process $\{B_n(t), t \geq 0\}$ is not continuous we define the following,

$$B_n^{(c)}(t) = B_n(t) - (nt - \lfloor nt \rfloor) \frac{Y_{\lfloor nt \rfloor}}{\sqrt{n}}$$

which is a continuous process and converges in law to a process $(B_t)_{t \geq 0}$ having the same law as $\tilde{B}$, and thus satisfies all properties of a Brownian Motion.

From this proposition we can thus see that by re scaling a random walk we retrieve a process satisfying the properties of the Brownian Motion.
2.4 The Cox-Ross-Rubinstein model

Now let us briefly introduce the Cox-Ross-Rubinstein model. Let \((Z_k)_{k \geq 0}\) be a sequence of i.i.d random variables such that \(\mathbb{P}[Z_k = 1] = \mathbb{P}[Z_k = -1] = \frac{1}{2}\). Let \(T > 0\) be some final time and \((b_n, \sigma_n)_{n \geq 1}\) the sequence defined by,

\[
b_n = b \frac{T}{n} \quad \text{and} \quad \sigma_n = \sigma \sqrt{\frac{T}{n}}
\]

where \(b, \sigma > 0\) are given.

In the Cox-Ross Rubinstein model we consider the price process of a single risky asset \(S^n = \{S^n_k, k = 0, \ldots, n\}\) to be,

\[
S^n_k = S^n_0 \exp(kb_n + \sigma_n \sum_{i=1}^k Z_i)
\]

with \(S^n_0\) given.

In fact this model, also called the binomial model, was first proposed by William Shape in 1978, and formalized by Cox, Ross and Rubinstein in 1979. It is a common numerical method that is used for the pricing of financial assets. In the following we will be using a simpler model inspired by the above. Where the main difference is that the price process of our risky asset will be,

\[
S^n_k = S^n_0 + \sum_{i=1}^k c_n + \sigma_n \sum_{i=1}^k Z_i
\]

with \((c_n)_n\) a different deterministic sequence that we will later define.
3 Problem

Now that we have defined the required notions, it is possible to proceed to motive, define and tackle the problem.

3.1 Motivation

Once again, in this problem, we take the role of the market maker. The goal is to maximize the profit the bid-ask spread generates. This is not a straightforward problem because increasing this spread does not necessarily increase the profit. Indeed, increasing the gap between the ask and bid prices makes it so that we gain more per transaction. But in doing so, we would attract fewer clients to engage with said offers, and thus decreasing the number of transactions. In the following, we create a model that takes into account these two phenomena to find the optimal choice of ask and bid prices.

3.2 Formalization

Let us consider a stock $S$ from a time 0 to $T \in \mathbb{R}^+$ given. In order to better model reality, we have decided to take a discrete approach to the problem in contrast with the continuous approach taken by the paper [2] on which this research was inspired. Let $(S_i)_{0 \leq i \leq N}$ the prices of this stock in times $0, \Delta t, 2\Delta t, \ldots, N\Delta t$, with $\Delta t := \frac{T}{N}, N \in \mathbb{N}$. That is for the sake of simplicity $S_i$ is the price of $S$ at time $i\Delta t$. We define the mathematical objects required for the problem.

1. At time 0 the price $S_0$ is given, and we initialize the number of bid and ask orders as $\Delta N_0^a = \Delta N_0^b = 0$. The variances at each time $\sigma_0, \ldots, \sigma_N$ are given.

We will be working with a zero-intelligence model, which means we can model the arrivals of the bid and ask models as random variables.
2. $\Delta N^a_1 \sim \text{Poi}(\lambda^a_0(p^a_1))$ the number of ask orders from time 0 to time 1, with $\lambda^a_0(p^a_1) = Ce^{-\frac{(p^a_1 - S_0)}{\sigma_0\sqrt{\Delta t}}}, C \in \mathbb{R}$ given. We can write $\delta^a_0 = p^a_1 - S_0$, we call it the ask spread.

3. $\Delta N^b_1 \sim \text{Poi}(\lambda^b_0(p^b_1))$ the number of bid orders from time 0 to time 1, with $\lambda^b_0(p^b_1) = Ce^{-\frac{(S_0 - p^b_1)}{\sigma_0\sqrt{\Delta t}}}, C \in \mathbb{R}$ given as before. We can write $\delta^b_0 = S_0 - p^b_1$, we call it the bid spread.

We can note that defining $\Delta N^a_1$ and $\Delta N^b_1$ in this way allows us to model the effect the prices we choose for the financial asset will have on the incoming ask and bid orders. That is as we said before increasing any of the spreads decreases the number of orders and vice-versa.

4. We then define for $1 \leq i \leq N - 1$,

$$\Delta N^a_{i+1} \sim \text{Poi}(\lambda^a_i(p^a_{i+1})) \text{ and } \Delta N^b_{i+1} \sim \text{Poi}(\lambda^b_i(p^b_{i+1}))$$

the ask and bid orders at time $i$. With,

$$\lambda^b_i(p^b_{i+1}) = Ce^{-\frac{\delta^b_i}{\sigma_i\sqrt{\Delta t}}} \text{ and } \lambda^a_i(p^a_{i+1}) = Ce^{-\frac{\delta^a_i}{\sigma_i\sqrt{\Delta t}}}$$

and the spreads given by,

$$\delta^b_i = S_i - p^b_{i+1} \text{ and } \delta^a_i = p^a_{i+1} - S_i$$

5. We define the market impact at each step $i$ as follows,

$$M_i(p^a_i, p^b_i) := \mathbb{E}[p^a_i \Delta N^a_i - p^b_i \Delta N^b_i | \mathcal{F}_{i-1}]$$

with $M_0 = 0$, and $\mathcal{F}_i$ is the natural filtration of $\Delta N^b_i, \Delta N^a_i, S_i$ for each
As we can see in the following, the Market Impact models the change induced by the prices we set at time $i$ on the price of the stock at time $i+1$. This is why it is important to consider the conditional expectation with respect to the filtration $\mathcal{F}$.

6. We define recursively the price of the stock at each step,

$$S_{i+1} = S_{i} + M_{i}(p_{i}^{a}, p_{i}^{b}) + \sigma_{i} \sqrt{\Delta t}Z_{i}$$

with $\mathbb{P}[Z_{i} = 1] = \mathbb{P}[Z_{i} = -1] = \frac{1}{2}$ and $Z_{i}$ independent of $S_{i}$ for each $0 \leq i \leq N$.

With all of these objects defined we can properly state our problem.

**Problem** Solve

$$\sup_{(p_{a}, p_{b})} \mathbb{E}\left[ \sum_{i=1}^{N} p_{i}^{a} \Delta N_{i}^{a} - p_{i}^{b} \Delta N_{i}^{b} + S_{N}Q_{N} \right]$$

with $Q_{N} = N_{N}^{b} - N_{N}^{a} = \sum_{i=1}^{N} \Delta N_{i}^{b} - \sum_{i=1}^{N} \Delta N_{i}^{a}$, $p^{a} = (p_{1}^{a}, ..., p_{N}^{a})$ and $p^{b} = (p_{1}^{b}, ..., p_{N}^{b})$

**3.3 Intuition of the problem**

Before tackling the problem let us motivate the choice we made for each mathematical object as well as give some intuition on the most important ones. First of all, the choice of the dynamics of the prices of the stock at each time was inspired by the Cox-Ross-Rubinstein model previously introduced. Secondly, the Poisson distribution expresses the probability that a given number of events takes place in a fixed interval of time. And as said
before, by taking into account the effect of the spreads on the parameter of the Poisson random variables we obtain a solid model for the arrival of the orders.

Next, inside the expectation operator we have two important quantities. The profit and loss,

\[ \sum_{i=1}^{N} p^a_i \Delta N^a_i - p^b_i \Delta N^b_i \]

that represent the quantity we gain, or lose, by engaging with each ask offer with a price \( p^a_i \) and each bid offer at a price \( p^b_i \) at the \( i \)-th time.

Then the next term

\[ S_N Q_N \]

allows us to model the risk we take by engaging with all bid and ask offers. Indeed, \( Q_N := N^b_N - N^a_N \), represents the difference of how many shares of the stock we hold at the last time, since it is the difference between the bid and ask offers we engaged with up to time \( N \). Note that this quantity can be negative (debt) or positive (surplus). In the case where this quantity is negative it can be interpreted as, owing \( Q_N \) units worth \( S_N \). The case where this quantity is positive is analogous. We call \( S_N Q_N \) the inventory risk.

We can therefore define the total profit as the sum of these two terms, as it would globally take into account the profit we made by buying and selling the stock, plus the risk we take by engaging with all offers.

Knowing this our problem can now be interpreted as choosing the correct bid and ask prices at each time \( i \) in order to maximize the expected total profit.
4 Theoretical work

4.1 1-period problem

It is usually interesting to study the simple cases of problems of this nature. Let us in the following consider that $N = 1$ and solve for the optimal values $p^a_1$ and $p^b_1$. In other words let us solve,

$$V_0 = sup(p^a_1, p^b_1) E[p^a_1 \Delta N^a_1 - p^b_1 \Delta N^b_1 + S_1 Q_1]$$

We have that,

$$V_0 = sup(p^a_1, p^b_1) \left\{ p^a_1 E[\Delta N^a_1] - p^b_1 E[\Delta N^b_1] + E[S_1 Q_1] \right\}$$

As $S_0$ is given we have that, $E[\Delta N^a_1] = \lambda^a_0(p^a_1)$ and $E[\Delta N^b_1] = \lambda^b_0(p^b_1)$. By definition of $S_1$ and $Q_1$ we have,

$$E[S_1 Q_1] = E[(S_0 + \sigma_0 \sqrt{t} Z_0)(\Delta N^b_1 - \Delta N^a_1)]$$

as $M_0 = 0$. Moreover, $E[Z_0] = 0$ and that $Z_0$ and $S_0$ are independent, so

$$E[S_1 Q_1] = S_0(\lambda^b_0 - \lambda^a_0)$$

Thus we finally get,

$$V_0 = sup(p^a_1, p^b_1) \left\{ \lambda^a_0(p^a_1 - S_0) + \lambda^b_0(S_0 - p^b_1) \right\}$$

$$V_0 = sup(p^a_1, p^b_1) \left\{ \lambda^a_0 \delta^a_0 + \lambda^b_0 \delta^b_0 \right\}$$
Let us note \( U(p_a^1, p_b^1) = \lambda_0^a \delta_0^a + \lambda_0^b \delta_0^b \) Knowing that,

\[
\begin{align*}
\partial_{p_a^1} \lambda_0^a(p_a^1) &= \frac{-\lambda_0^a}{\sigma_0 \sqrt{\Delta t}} & \partial_{p_a^1} \delta_0^a &= 1 \\
\partial_{p_b^1} \lambda_0^b(p_b^1) &= \frac{\lambda_0^b}{\sigma_0 \sqrt{\Delta t}} & \partial_{p_b^1} \delta_0^b &= -1
\end{align*}
\]

we obtain that,

\[
\nabla U(p_a^1, p_b^1) = \begin{pmatrix}
\frac{-\lambda_0^a}{\sigma_0 \sqrt{\Delta t}} (\delta_0^a - \sigma_0 \sqrt{\Delta t}) \\
\frac{\lambda_0^b}{\sigma_0 \sqrt{\Delta t}} (\delta_0^b - \sigma_0 \sqrt{\Delta t})
\end{pmatrix}
\]

Therefore the critical values are obtained for the following values of the spread,

\[
\begin{align*}
\delta_0^a &= \sigma_0 \sqrt{\Delta t} \\
\delta_0^b &= \sigma_0 \sqrt{\Delta t}
\end{align*}
\]

Which induces the following optimal values,

\[
\begin{align*}
p_a^1 &= S_0 + \sigma_0 \sqrt{\Delta t} \\
p_b^1 &= S_0 - \sigma_0 \sqrt{\Delta t}
\end{align*}
\]

Finally By plugging everything into \( V_0 \) we obtain,

\[
V_0 = \frac{2C}{e^{\sigma_0 \sqrt{\Delta t}}}
\]

### 4.2 2-period problem

The result obtained for the 1-period problem being quite promising, we study now the next step in order to look for interesting patterns. That is we want
to solve the problem for $N = 2$

$$V_0 = \sup_{(p_1^a, p_2^a, p_1^b, p_2^b)} \mathbb{E} \left[ p_1^a \Delta N_1^a - p_1^b \Delta N_1^b + p_2^a \Delta N_2^a - p_2^b \Delta N_2^b + S_2 Q_2 \right]$$

Consider the following quantities,

$$V_2 = S_2 Q_2$$
$$V_1 = \sup_{(p_2^a, p_2^b)} \left\{ M_2 + \mathbb{E}[V_2|\mathcal{F}_1] \right\}$$
$$V_0' = \sup_{(p_1^a, p_1^b)} \left\{ M_1 + \mathbb{E}[V_1] \right\}$$

Next we admit that the $\sup$ and expectation can be exchanged, indeed this fact is sometimes referred to as the Bellman’s Principle of optimality. Plugging everything in $V_0'$ we obtain,

$$V_0' = \sup_{(p_1^a, p_1^b)} \left\{ M_1 + \mathbb{E} \left[ \sup_{(p_2^a, p_2^b)} \left\{ M_2 + \mathbb{E}[S_2 Q_2|\mathcal{F}_1] \right\} \right] \right\}$$
$$V_0' = \sup_{(p_1^a, p_1^b)} \left\{ M_1 + \mathbb{E} \left[ \mathbb{E}[S_2 Q_2|\mathcal{F}_1] \right] \right\} \text{ by exchanging the } \sup \text{ and } \mathbb{E}$$
$$V_0' = \sup_{(p_1^a, p_1^b)} \left\{ \sup_{(p_2^a, p_2^b)} \left\{ M_1 + \mathbb{E}[M_2] + \mathbb{E} \left[ \mathbb{E}[S_2 Q_2|\mathcal{F}_1] \right] \right\} \right\} \text{ as } M_1 \text{ is independent of } (p_2^a, p_2^b)$$
$$V_0' = \sup_{(p_1^a, p_1^b, p_2^a, p_2^b)} \left\{ M_1 + \mathbb{E}[M_2] + \mathbb{E}[S_2 Q_2] \right\} \text{ by the law of total expectation}$$

Finally as $\mathbb{E}[M_2] = \mathbb{E}[p_2^a \Delta N_2^a - p_2^b \Delta N_2^b]$ by the law of total expectation, and the definition of $M_2$. We finally identify $V_0 = V_0'$. This allows us to proceed in a modular way and optimize the pairs of $(p_i^a, p_i^b)_{i \in \{1,2\}}$ one step at a time.

Now consider,

$$V_2 = S_2 Q_2 = (S_1 + M_1 + \sigma_1 \sqrt{\Delta t} Z_1)(\Delta N_2^a - \Delta N_2^b + Q_1)$$
then denoting $p_2 = (p_a^2, p_b^2),

\begin{align*}
V_1 &= \sup_{(p_2)} \left\{ M_2 + \mathbb{E}[V_2 | F_1] \right\} \\
&= \sup_{(p_2)} \left\{ p_a^2 \lambda_1^a - p_b^2 \lambda_1^b + \mathbb{E}[(S_1 + M_1 + \sigma_1 \sqrt{\Delta t} Z_1)(\Delta N_2^a - \Delta N_2^b + Q_1) | F_1] \right\} \\
&= \sup_{(p_2)} \left\{ p_a^2 \lambda_1^a - p_b^2 \lambda_1^b + (S_1 + M_1)(Q_1 + \lambda_1^b - \lambda_1^a) \right\} \text{ as } S_1, Q_1 \text{ and } M_1 \text{ are } F_1 \text{ measurable}
\end{align*}

\begin{align*}
V_1 &= \sup_{(p_2)} \left\{ (p_a^2 - S_1 - M_1) \lambda_1^a + (S_1 + M_1 - p_b^2) \lambda_1^b \right\} + (S_1 + M_1)Q_1
\end{align*}

Denote $U_1(p_a^2, p_b^2) = (p_a^2 - S_1 - M_1)) \lambda_1^a + (S_1 + M_1 - p_b^2) \lambda_1^b$, let us optimize it. We have

\begin{align*}
\nabla U_1(p_a^2, p_b^2) &= \begin{pmatrix}
-\lambda_1^a \frac{\sigma_1 \sqrt{\Delta t}}{\sigma_1 \sqrt{\Delta t}}(p_a^2 - S_1 - M_1 - \sigma_1 \sqrt{\Delta t}) \\
\lambda_1^b \frac{\sigma_1 \sqrt{\Delta t}}{\sigma_1 \sqrt{\Delta t}}(-p_b^2 + S_1 + M_1 - \sigma_1 \sqrt{\Delta t})
\end{pmatrix}
\end{align*}

Thus, solving $\nabla U_1(p_a^2, p_b^2) = (0, 0)$ we get

\begin{align*}
p_a^2 &= S_1 + M_1 + \sigma_1 \sqrt{\Delta t} \\
p_b^2 &= S_1 + M_1 - \sigma_1 \sqrt{\Delta t}
\end{align*}

We observe that we obtain a similar result as in the 1-period case. Plugging these values back into the expression we can express the $V_1$ explicitly,

\begin{align*}
V_1 = (S_1 + M_1)Q_1 + 2e^{-1}C\sigma_1 \sqrt{\Delta t} \cosh \left( \frac{M_1}{\sigma_1 \sqrt{\Delta t}} \right)
\end{align*}

Denoting $\beta = 2e^{-1}C$ and $p_1 = (p_1^a, p_1^b)$, we get finally,
\[ V_0 = \sup_{p_1} \left\{ M_1 + \beta \sigma_1 \sqrt{\Delta t} \cosh \left( \frac{M_1}{\sigma_1 \sqrt{\Delta t}} \right) + \mathbb{E}[(S_1 + M_1)Q_1] \right\} \]

\[ V_0 = \sup_{p_1} \left\{ M_1 + \beta \sigma_1 \sqrt{\Delta t} \cosh \left( \frac{M_1}{\sigma_1 \sqrt{\Delta t}} \right) + (S_0 + M_1)(\lambda_0^b - \lambda_0^a) \right\} \]

Denote \( U_0 = M_1 + \beta \sigma_1 \sqrt{\Delta t} \cosh \left( \frac{M_1}{\sigma_1 \sqrt{\Delta t}} \right) + (S_0 + M_1)(\lambda_0^b - \lambda_0^a) \), and note that

\[ \partial_{p_2^a} M_1 = \frac{\lambda_0^a}{\sigma_0 \sqrt{\Delta t}} (\sigma_0 \sqrt{\Delta t} - p_1^a) \]

\[ \partial_{p_1^b} M_1 = \frac{-\lambda_0^b}{\sigma_0 \sqrt{\Delta t}} (\sigma_0 \sqrt{\Delta t} + p_1^b) \]

The gradient of \( U_0 \) is

\[ \nabla U_0(p_1^a, p_1^b) = \left( \frac{\lambda_0^a}{\sigma_0 \sqrt{\Delta t}} \left( S_0 + M_1 + \sigma_0 \sqrt{\Delta t} - p_1^a \right) \left( 1 + \beta \sinh \left( \frac{M_1}{\sigma_1 \sqrt{\Delta t}} \right) + \lambda_0^b - \lambda_0^a \right) \right) \]

\[ \left( \frac{-\lambda_0^b}{\sigma_0 \sqrt{\Delta t}} \left( S_0 + M_1 - \sigma_0 \sqrt{\Delta t} + p_1^b \right) \left( 1 + \beta \sinh \left( \frac{M_1}{\sigma_1 \sqrt{\Delta t}} \right) + \lambda_0^b - \lambda_0^a \right) \right) \]

We notice that due to the combination of nested exponential (inside the sinh function), exponential and linear functions we cannot compute explicitly the critical points. Nevertheless, this constitutes a necessary condition for an optimal point, thus we can identify at least identify them. (see notebook)

4.3 Derivation of the recursive identity

We have thus seen that finding an explicit formula to the problem might not be feasible. Moreover, computing the critical points becomes more convoluted as \( N \) increases. Nevertheless, in the above problem we have managed
to decompose the optimization in steps. That is why in the following we will be tackling the problem with a Dynamic Programming approach. Let us start by denoting

\[ V_n = \sup_{(p_i)_{i>n}} \left\{ \mathbb{E} \left[ \sum_{i=n+1}^{N} p_i^a \Delta N_i^a - p_i^b \Delta N_i^b + S_N Q_N \bigg| \mathcal{F}_n \right] \right\} \]

with \( \mathcal{F}_n \) the natural filtration of \( \Delta N_i^a, \Delta N_i^b, Z_n \). We therefore can retrieve our original problem as follows.

\[ V_0 = \sup_{(p_i)_{N>0}} \left\{ \mathbb{E} \left[ \sum_{i=1}^{N} p_i^a \Delta N_i^a - p_i^b \Delta N_i^b + S_N Q_N \right] \right\} \]

since \( \mathcal{F}_0 \) is the trivial filtration.

In order to apply Dynamic programming we seek to create a recursive relation. We start with,

\[ V_N = S_N Q_N \]

Then for any \( n < N \) we have,

\[ V_n = \sup_{(p_i)_{i>n}} \left\{ \mathbb{E} \left[ \sum_{i=n+1}^{N} p_i^a \Delta N_i^a - p_i^b \Delta N_i^b + S_N Q_N \bigg| \mathcal{F}_n \right] \right\} \]

\[ = \sup_{(p_i)_{i>n}} \left\{ M_{n+1} + \mathbb{E} \left[ \sum_{i=n+2}^{N} p_i^a \Delta N_i^a - p_i^b \Delta N_i^b + S_N Q_N \bigg| \mathcal{F}_n \right] \right\} \]

\[ = \sup_{(p_i)_{i>n}} \left\{ M_{n+1} + \mathbb{E} \left[ \sum_{i=n+2}^{N} p_i^a \Delta N_i^a - p_i^b \Delta N_i^b + S_N Q_N \bigg| \mathcal{F}_{n+1} \bigg| \mathcal{F}_n \right] \right\} \]

By the Tower property as \( \mathcal{F}_n \subset \mathcal{F}_{n+1} \). Next using Bellman’s Principle of optimality we can exchange the sup and \( \mathbb{E} \).

Denote \( \alpha := \mathbb{E} \left[ \sum_{i=n+2}^{N} p_i^a \Delta N_i^a - p_i^b \Delta N_i^b + S_N Q_N \bigg| \mathcal{F}_{n+1} \right] \), so
\[ V_n = \sup_{(p_i)_{i>n}} \left\{ M_{n+1} + \mathbb{E} \left[ \alpha \mid \mathcal{F}_n \right] \right\} \]
\[ = \sup_{p_{n+1}} \left\{ M_{n+1} + \sup_{(p_i)_{i>n+1}} \left\{ \mathbb{E} \left[ \alpha \mid \mathcal{F}_n \right] \right\} \right\} \]

since \( M_{n+1} \) depends only on \( p_{n+1} \). Moreover exchanging the sup and the expectation.

\[ = \sup_{(p_{n+1})} \left\{ M_{n+1} + \mathbb{E} \left[ \sup_{(p_i)_{i>n+1}} \alpha \right] \mathcal{F}_n \right\} \]

We see that \( V_{n+1} = \sup_{(p_i)_{i>n+1}} \alpha \) by definition. Therefore,

\[ V_n = \sup_{p_{n+1}} \left\{ M_{n+1} + \mathbb{E} \left[ V_{n+1} \mid \mathcal{F}_n \right] \right\} \]

We thus get the following recursive relation.

\[
\begin{align*}
V_N &= S_N Q_N \\
V_n &= \sup_{p_{n+1}} \left\{ M_{n+1} + \mathbb{E} \left[ V_{n+1} \mid \mathcal{F}_n \right] \right\}
\end{align*}
\]

5 **Numerical Solution**

With the recursive identity we have retrieved above we can start applying the Dynamic Programming principle. That is at each step we will assume we know \( V_{n+1} \) in order to compute \( V_n \). To do so it is interesting to consider \( V_{n+1} \) as a function of \( S_n, Q_n \). We have,

\[ V_n(S_n, Q_n) = \sup_{(p_{n+1})} \left\{ M_{n+1} + \mathbb{E} \left[ V_{n+1}(S_{n+1}, Q_{n+1}) \mid \mathcal{F}_n \right] \right\} \]
By definition,
\[
\begin{aligned}
S_{n+1} &= S_n + M_n + \sigma_n \Delta t Z_n \\
Q_{n+1} &= Q_n + \Delta N^b_n - \Delta N^a_n
\end{aligned}
\]

So since $S_n, Q_n$ are $\mathcal{F}_n$-measurable, and $\Delta N^b_n, \Delta N^a_n, Z_n$ are all independent of $\mathcal{F}_n$

\[
V_n(S_n, Q_n) = \sup_{(p_{n+1})} \left\{ M_{n+1} + \mathbb{E} \left[ V_{n+1}(S_n + M_n + \sigma_n \Delta t Z_n, Q_n + \Delta N^b_n - \Delta N^a_n) \right] \right\}
\]

Thus taking the expectation with respect to the first and second variables we get,

\[
V_n(S_n, Q_n) = \sup_{(p_{n+1})} \left\{ M_{n+1} + \frac{1}{2} \sum_{k_a=0}^{\infty} \sum_{k_b=0}^{\infty} (V^+ + V^-) e^{-\lambda_n b} \frac{(\lambda_n b)^{k_b}}{k_b!} e^{-\lambda_n a} \frac{(\lambda_n a)^{k_a}}{k_a!} \right\}
\]

with $V^+ = V_{n+1}(S_n + M_n + \sigma_n \Delta t, Q_n + k_b - k_a),$ and $V^- = V_{n+1}(S_n + M_n - \sigma_n \Delta t, Q_n + k_b - k_a)$.

Thus with this relation we can implement a numerical solution interpolating $V_{n+1}$. Indeed our strategy will be the following.

1. At step $N$ we know the exact structure of the function $V_N(S_N, Q_N) = S_N Q_N$ thus we use a grid to compute the function in the necessary points in order to interpolate it.

2. Then for $n < N$, as we have an interpolated function $V_{n+1}$ we can compute $V_n$ using the above formula, i.e we will compute the inner part of the formula for different values of $p_{n+1}$ and keep the maximal value. In doing so we will create a mapping from the the values of the grid $(S_n, Q_n)$ to the corresponding optimal prices $p_{n+1}$. After knowing $V_n$ in several points we can interpolate it and move on to $n - 1$. 

19
5.1 Defining the grid

Let us now explain how to construct the adequate grids for each variable. First, in order to create the grid of points for the first variable of $V_n, S_n$ we use the binomial model on which this random variable is based upon (see Figure 1). It is easy to see that this tree is recombinant, which means "going up and then down" is the same as "going down and then up" in the tree. This ultimately means that there are $n$ possible values at step $n$. They are the nodes $n$ edges away from the origin, that is $G_n = \{\tilde{S}_{n,i}, 0 \leq i \leq n\}$.

We remark that creating this grid we ignore the market impact, that is because as $N \to \infty$ the market impact goes to $0$, so the accuracy of our binomial model increases. Moreover, this model only fails in the case when $S_n + M_n \notin [\inf G_n, \sup G_n]$ because of the interpolation (That is because the interpolation technique we will use in the following fails outside of the domain given), so our grid is adequate even for small values of $N$.

Next, in order to make the grid for $Q_n := N_N^b - N_N^a$ it is just a matter to choose enough values in $\mathbb{N}$ symmetrically and with the corresponding grid at step $n - 1$ included in the grid at step $n$. This is just due to the fact that $Q_n$
is a sum and difference of Poisson random variables. We have nevertheless to keep in mind that we need a significant amount of points for the interpolation.

Finally we will also need a grid for $p_n$ at each step to perform the optimization. Recalling that $\delta_n^a = p_{n+1}^a - S_n$ and $\delta_n^b = S_n - p_{n+1}^b$, it is clear that we can also optimize with respect to these values. Indeed it is easier and more intuitive to proceed in this way. We start by choosing a tick size, i.e the smallest possible quantity. For example we could choose 0.01$ as in real life the smallest possible currency are cents. Then we choose biggest spread we are willing to choose, i.e 2$. Therefore the grid we will iterate over is all the possible values in this range, that is 0.01$, 0.02$, \ldots, 2.00$.

5.2 Results

We proceed to implement the solution in the Jupyter Notebook attached to this report.

First of all to test our model we need to choose numerical values for the parameters defined above. In the following we will set $C = 1$, this value is the amplitude of our Poisson parameters, the final time $T = 1$, the tick size $= 0.01$, we set the biggest possible spread to be 100 units and $S_0 = 100$. Finally inspired by the volatility of the apple stock (APPl [4]) we set our $\sigma_i = 0.3$ for each $0 \geq i \geq N$ as we can see that this number is quite stable.

In our Notebook, we start by revisiting the 1 and 2-period problems, and we confirm that our algorithm agrees with the theoretical facts we have stated above. It is also where we present the format we have decided to use for the mapping $P$ that takes as inputs values of the price of the stock $S$ and the difference of bid and ask arrival orders $Q$ and outputs the optimal bid and ask prices $p^a, p^b$ that we can set. An empirical comment we can make, is that at least for small values of $N$ this mapping seems to me independent of $Q$, which is in agreement once again with our theoretical results.

After treating the simple cases we proceed to test our solution for bigger
values of $N$. As we do not know the optimal values for this case we cannot directly confirm that the results obtained are adequate. Therefore we perform a statistical analysis of our results and obtain the following.

First of all we can follow a random path, i.e simulate $(Z_t)$ (see notebook), and plot the corresponding optimal prices we obtain at each step. Here are some examples. From Figure 2 we see that as we expected as we are using the re-scaled random walk in our model, these graphs look like zoomed in Brownian Motions. More specifically since they do not start from 0, and by the nature of our model we can say that as $\Delta t \to 0$ these prices go to an arithmetic Brownian Motion.

Next we proceed to study the expectation and the variance of these prices, recalling that as we saw in the 2-period problem they are random processes. We see the graphs for these two operators in Figure 3. First of all the Expectation graph shows a slight abnormal behaviour at the right extreme, a sudden drop. This could be indicative of a lost of precision as we approach $N$.

Then for the variance, we can see that it increases linearly with time.
Recalling the solution for the 2-period problem

\[ p_a^2 = S_1 + M_1 + \sigma_1 \sqrt{\Delta t} \]
\[ p_b^2 = S_1 + M_1 - \sigma_1 \sqrt{\Delta t} \]

we can see that applying the Variance on both values yields \( \sigma_0 \sqrt{\Delta t} \). Indeed, even though \( M_i \) is a random variable seen from time 0 it can be considered as a constant seen from time \( i - 1 \), which is the time we are in in when it comes to finding the optimal \( p_a^i, p_b^i \). The linearity of the graph gives us a hint. Perhaps we might have

\[ p_a^i = S_{i-1} + c_1 \]
\[ p_b^i = S_{i-1} + c_2 \]

For some constants \( c_1, c_2 \in \mathbb{R} \).

As this would yield

\[ Var(p_a^i) = Var(p_b^i) = \sqrt{\Delta t} \sum_{j=0}^{i-2} \sigma_j \]

Finally after seeing the results it is interesting to discuss the work that can be done to further our understanding of the problem. It is also natural to proceed to identify the limitations of our model and to propose solutions to them.
6 Future Work

In conclusion, in this paper we have have managed to motivate and formalize the problem involving the dynamics of a stock prices and its relation with the bid-ask spread, as well as to provide a numerical solution to it. Now as this research was done in a constrained time interval it seems adequate to describe the limitations of our solution as well as introduce some possible next steps that could be taken to further explore this problem.

Firstly, regarding the limitations. The biggest one is the fact that our algorithm is significantly slow. It takes about 30 minutes to run the solution for \( N = 40 \), with the first steps being much slower than the rest, meaning that the time this program takes to execute increases exponentially.

Thus a natural next step is to continue optimizing our code, in order to be able to solve the problem for bigger values of \( N \), and thus obtain a more applicable model. And consequently put it to the test with data extracted from real life.

Secondly, as we said in section 2.4, we used a simpler model inspired by the Cox-Ross-Rubinstein model in order to solve our problem. As it is a common practice in mathematical finance, a logical next step would be to implement the actual Cox-Ross-Rubinstein model to solve our problem.

Finally, using the statistical analysis from above we could continue exploring the theory behind the problem in order to find more interesting patterns to aid our algorithm.
References


