Fire Sale in High Frequency Order Book

Dissertation Submitted In Partial Fulfillment Of The Requirements For The Degree Of

Master of Technology
in
Computer Science

by

Debam Kumar Das
[Roll No: CS1805]

Under the Guidance of

Dr. Diganta Mukherjee
Professor
Sampling and Official Statistics Unit(SOSU)

Indian Statistical Institute
Kolkata-700108, India

CERTIFICATE

This is to certify that the dissertation entitled **"Fire Sale in High Frequency Order Book"** submitted by **Debam Kumar Das** to Indian Statistical Institute, Kolkata, in partial fulfilment for the award of the degree of **Master of Technology in Computer Science** is a *bona fide* record of work carried out by him under my supervision and guidance. The dissertation has fulfilled all the requirements as per the regulations of this institute and, in my opinion, has reached the standard needed for submission.

Dr. Diganta Mukherjee
Professor,
Sampling and Official Statistics Unit,
Indian Statistical Institute,
Kolkata-700108, India.

Abstract: This work attempts to come up with an optimal strategy that a market-maker could adopt in times of a fire sale at the high frequency level. We use a modern version of the self financing equation applied to a high frequency order book to model the market dynamics. Using the model, we setup an optimal stochastic control problem, the solution to which is the optimal strategy proposed.

1 Introduction:

1.1 Fire sale:

A fire sale [see 6] is essentially a forced sale of an asset at a dislocated price. The asset sale is forced in the sense that the seller cannot pay creditors without selling assets. The price is dislocated because the highest potential bidders are typically involved in a similar activity as the seller, and are therefore themselves indebted and cannot borrow more to buy the asset. Indeed, rather than bidding for the asset, they might be selling similar assets themselves. Assets are then bought by nonspecialists who, knowing that they have less expertise with the assets in question, are only willing to buy at valuations that are much lower. See figure 1 for visualization.

1.2 Market maker:

Traditionally market-makers used to be large banks or financial institutions. The role of market-maker is to ensure that there's enough liquidity in the market. They make sure that any buyer or seller don't have much hassle in finding a giver and taker of the corresponding asset. They continually quote the 'bid' and 'ask' prices at which they want to buy and sell. Anyone wishing to sell or buy can match their prices and participate in a transaction.

Market-maker profit from the b̈id-ask spreadẅhich is the difference between the ask and the bid price they quote, i.e. the difference in the price at which they buy and sell the asset. The first issue faced by an market-maker when providing liquidity is that by accepting one side of a trade (say buying from someone who wants to sell), the market-maker will hold an asset for an uncertain period of time, the time it takes for another person to come to the market with a matching demand for liquidity (wanting to buy the asset the market-maker bought in the previous trade). During that time, the market-maker is exposed to the risk that the price moves against her (in our example, as she bought the asset, she is exposed to a price decline and hence having to sell the asset at a loss in the next trade). The market-maker has no intrinsic need or desire to hold any inventory, so she will only buy (sell) in anticipation of a subsequent sale (purchase)[see 2].

1.3 Market making in the high frequency domain:

In recent years, with the growth of electronic exchanges such as NASDAQ's Inet, anyone willing to submit limit orders in the system can effectively play the role of a market-maker. Indeed, the availability of high frequency data on the limit order book ensures a fair playing field where various agents can post limit orders at the prices they choose.

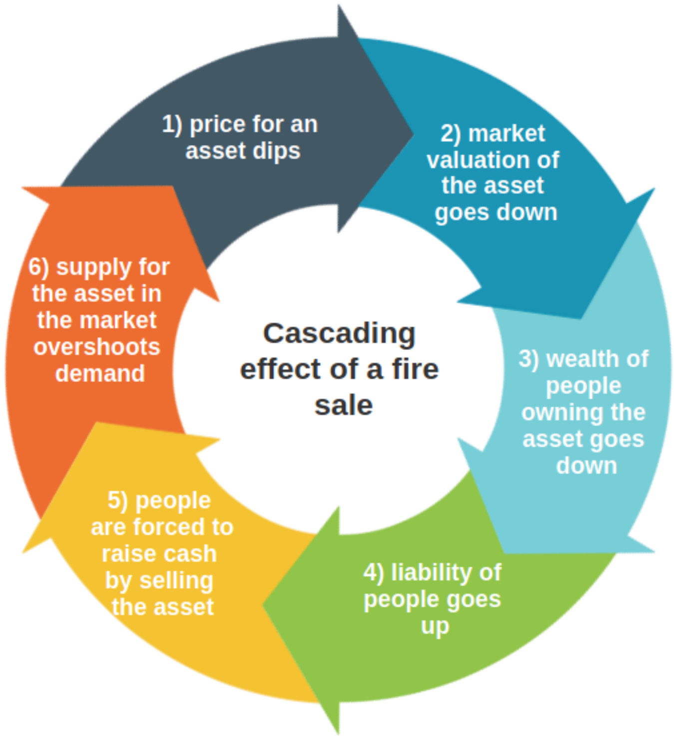

Figure 1: Cascading effect of a Fire Sale on market price

In this work, we plan to study the optimal submission strategies of bid and ask orders in a limit order book. Apart from the above mentioned inventory risk arising from the uncertainty in asset price another risk that a market-maker is exposed to is the asymmetric information risk arising from informed traders. Many trades originate not because someone needs cash and sells an asset, or has extra cash and wants to invest, but because one party has (or believes she has) better information about what the price is going to do than is reflected in current prices. A "fire-sale" is one such instance where a certain mass of informed traders are informed about the price-slash that's looming in the near future and enter the market to leverage their additional information.
From the point of view of a market-maker who continually participates in the market, she needs to be wary of such a situation and continue quoting the bid and ask prices. This work hopes to address the above problem from the point of view of a single market-maker.

1.4 Self financing equation and its evolution:

1.4.1 Standard self financing equation:

The self financing equation is a very old development. It has been used in the classical Black-Scholes option pricing and Merton portfolio theories[see 8]. Following is the most primitive version corresponding to the friction-less case:

$$dX_t = L_t dp_t \tag{1}$$

1.4.2 Almgren and Chris:

The works of Almgren and Chriss [see 1] builds up on it. These authors proposed a macroscopic model for the price impact and the change of wealth after a liquidity taker's decision. The model leads to a very tractable frame-work which was, and still is, used in many optimal execution studies. This framework can be summarised by the system:

$$\begin{aligned} dp_t &= f(l_t)dt + \sigma_t dWt \\ dL_t &= l_t dt \\ dX_t &= L_t dp_t - c(l_t)dt \end{aligned} \tag{2}$$

where $f()$ and $c()$ are positive functions.

The main advantage of this model is that price impact appears in a tractable fashion. Indeed, it comes through the drift $f(l_t)$ of the price process, which creates a positive correlation between traded volume and price. However, it constrains L_t to be a differentiable function of time, and as a result, the model parameters cannot be calibrated to market data directly, making the model difficult to test empirically. As per the empirical studies reported in the appendices, there is ample evidence supporting non-differentiable inventories. Moreover, certain trading strategies such as delta-hedging, latency arbitrage and statistical arbitrage naturally lead to inventory models with infinite variation. Finally, note that the use of limit orders is not covered by Almgren and Chriss.

1.4.3 Carmona and Webster:

The works of Carmona and Webster [see 4, 5] presents a mathematical framework for trading on a limit order book, including its associated transaction costs, and also proposes a continuous-time equations which generalises the self-financing relationships of frictionless markets. These equations naturally differentiate between trading via limit and via market orders, as they include a price impact or adverse selection constraint. Moreover they do not constraint the inventory process to be differentiable. Their works can be mathematically summarised by the following set of equations:

$$\begin{aligned} dp_t &= \mu_t dt + \sigma_t dW_t \\ dL_t &= b_t dt + l_t dW_t' \\ dX_t &= L_t dp_t \pm \int_{\mathbb{R}} c(y)\phi_{\eta_t}(y)dydt + d[L,p]_t \end{aligned} \tag{3}$$

where ϕ_σ is the density function of the Gaussian distribution with mean 0 and variance σ^2.

The above self-financing equations are bare-bones descriptions of the market as they merely provide an accountant perspective. Given a trader's inventory and the limit order book he or she trades on, these equations state that the accountant can track his or her wealth perfectly. [4] further models three phenomenons vividly observed in the high frequency realm, price-impact, price-recovery and adverse selection. Adverse selection here is the liquidity takers (participants who trade via market-orders) being more more knowledgeable than the liquidity providers. Because of which the price always moves against the liquidity providers.

$$d[L,p] < 0 \tag{4}$$

where L is the liquidity provider's inventory process and p is the price process. The above equation can also be perceived as an illustration of price impact, $\Delta L' = -\Delta L$ where the process L' is the liquidity taker's inventory process. For price recovery, the following constraints are put

$$|\Delta p| \leq s \tag{5}$$

where s_t is the process denoting the spread (difference between best ask and best bid prices) at time t. [5] further goes on to suggest an application of their proposed wealth equation in market-making. To make sure that price impact constraints are satisfied, they propose, at the microscopic level, a modified version of the Almgren and Chriss [1, see] model to relate the change in mid-price to the change in the aggregate inventory of the liquidity providers as

$$\Delta_n L = -\lambda_{n+1} \Delta_n p \tag{6}$$

for a $\mathscr{F}_{n+1}$ measurable, positive random variable λ_{n+1}. The choice of λ_{n+1} is made inspired by the framework proposed by the works of Avallaneda and Stoikov [3, 9, see].

1.4.4 Our work:

We plan to model and solve the optimal control problem of maximising a single market-maker's earning in a market with falling prices. Traditionally, this problem has been studied in a way which requires numerous assumption regarding different market micro structures all working in tandem [9, 3, see]. The novelty of our work is the fewer assumptions regarding the arrival of the different kinds of orders and their interaction with the order book which allows for greater versatility in our model and hence more accurate depiction of the real world. We work with a non-parametric system with respect to the distribution of the arrival and execution of a trade order.

1.5 Our work flow:

The above mentioned work has been carried out in the following steps:

- A price dynamics has been proposed corresponding to the fire sale environment and has been empirically searched for validation. We used historical top NIFTY performer's data corresponding to days with historical price decline.
- With the assumption of a single market-maker in the market, we model the interaction of a quoted buy and ask price of the market-maker with the entire market.
- Using the above two elements we design a stochastic optimal control problem where the control parameter is the bid-ask spread every second and the state variables are price, market-maker's wealth (which includes her inventory and cash process). Our objective function would be to maximise market-maker's terminal wealth. We have used Pontryagin's maximum principle [7, 10, see] to solve the optimal control problem.
- Once the optimal control is arrived at, using our price model we create synthetic price paths (possible realizations of the price process) and see the distribution of the expected market-maker's wealth.

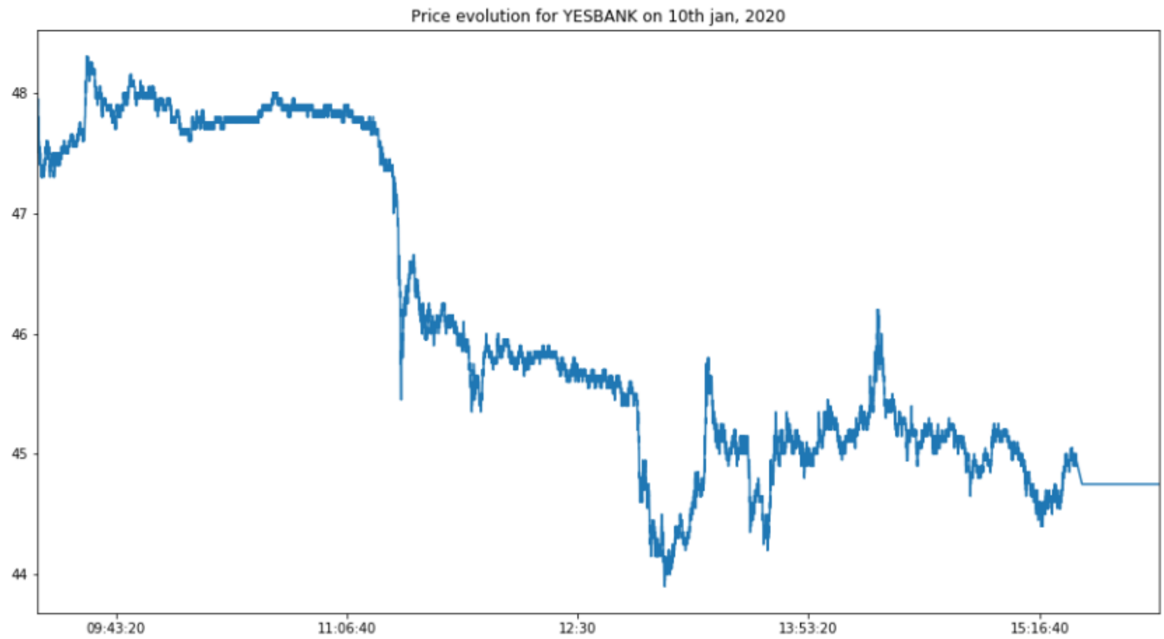

Figure 2: Price evolution for YESBANK on 10th Jan, 2020

2 Model Setup:

2.1 Price dynamics:

2.1.1 Formulation:

We model the price dynamics as:

$$dp_t = \mu_t dt + \sigma_t dW_t \tag{7}$$

The dp_t and dp_t are forward differentials and dW_t is a wiener's process. Our period of observation starts from time t_0 and ends at time t_f. The price at time t is given by p_t. We differentiate between the informed and the uninformed trader community in our model. The market-maker at any point of time need to be wary of the former community because they are a greater adversarial risk to her. We model the informed community with comparatively more precise knowledge of the future price movements. As a result of which the stochastic shocks in the price movements from the informed community will have a smaller magnitude. We go on to model stochastic component of the price movement in the following manner:

$$\sigma_t = \sqrt{f_{\frac{\sigma}{k}}(c_t)^2 + f_\sigma(1 - c_t)^2} \tag{8}$$

The function f represents the uncertainty brought in by a community. We expect the total uncertainty of the two communities to be dependent on their proportion. And alpha characterizes their preciseness.

$$f_\alpha(c_t) = \alpha\sqrt{c_t} \tag{9}$$

Here c_t is the proportion of the informed trader at time t. Therefore $1-c_t$ is the proportion of the uninformed trader. We go on to assume the two proportions remain constant through out our observation window.

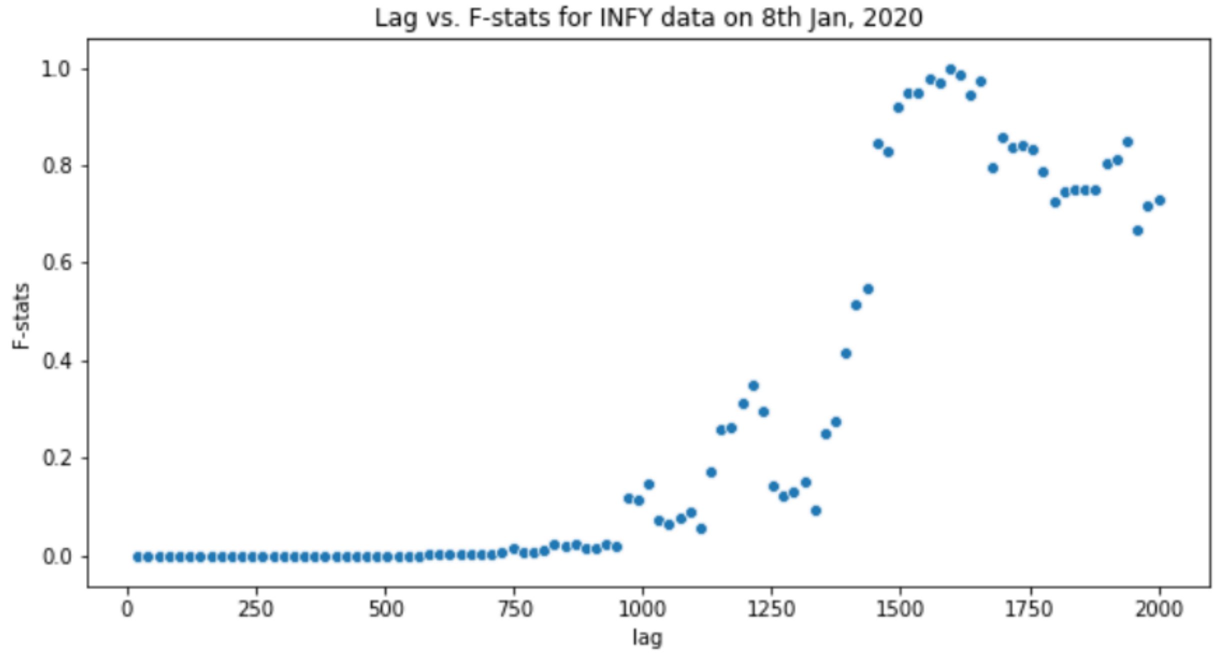

Figure 3: Lag vs. F-values on INFY logs for 8th Jan, 2020

We know that price during a fire trade is very volatile, and often depends majorly on sentiments. It is mostly the panic (specially among the uninformed traders) which decide the subsequent price path. To incorporate that idea into our model we model the drift as:

$$\begin{aligned} \mu_t &= \kappa(p_t - \phi(t)) \\ \phi(t) &= \min p_s \quad wheres \in [t-\delta, t] \end{aligned} \tag{10}$$

κ here is a constant which is to be determined by past data. And $\delta > 0$ is a lag window to be fine tuned as per the data.

2.1.2 Empirical evidence:

We regressed the Δp_t (the dependent variable) with $p_t - \phi(t)$ (the independent variable) without any intercept for several lags (δ) and chose a lag which had proper sign for κ (negative) and good significance. Data used is INFY 8th january, 2020 (see figure 2).Following is a plot of F-values vs. different lags (Figure 3). And the resulting quantile-quantile plot for the residues (5) plotted against standard normal distribution. The regression summary is in figure 4.

2.2 Order book shape:

Later when we try to model the interaction of the limit orders posted by the market-maker, we'll be required to have a rough idea how the order-book looks like at any point of time.More precisely, to model λ. $q_b(p)$ is the number of buy bids at price p. In this section we'll show the justification for an exponential shape of the order book on the buy and sell side.

```
                            OLS Regression Results
=======================================================================================
Dep. Variable:              del_price   R-squared (uncentered):                   0.001
Model:                            OLS   Adj. R-squared (uncentered):              0.000
Method:                 Least Squares   F-statistic:                              11.18
Date:                Wed, 08 Jul 2020   Prob (F-statistic):                    0.000829
Time:                        21:00:17   Log-Likelihood:                          22180.
No. Observations:               21615   AIC:                                 -4.436e+04
Df Residuals:                   21614   BIC:                                 -4.435e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
drift_min     -0.0016      0.000     -3.343      0.001      -0.003      -0.001
==============================================================================
Omnibus:                     1237.967   Durbin-Watson:                   2.569
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             5238.758
Skew:                           0.060   Prob(JB):                         0.00
Kurtosis:                       5.409   Cond. No.                         1.00
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 4: Price model estimation regression summary. Data: INFY 8th jan, 2020

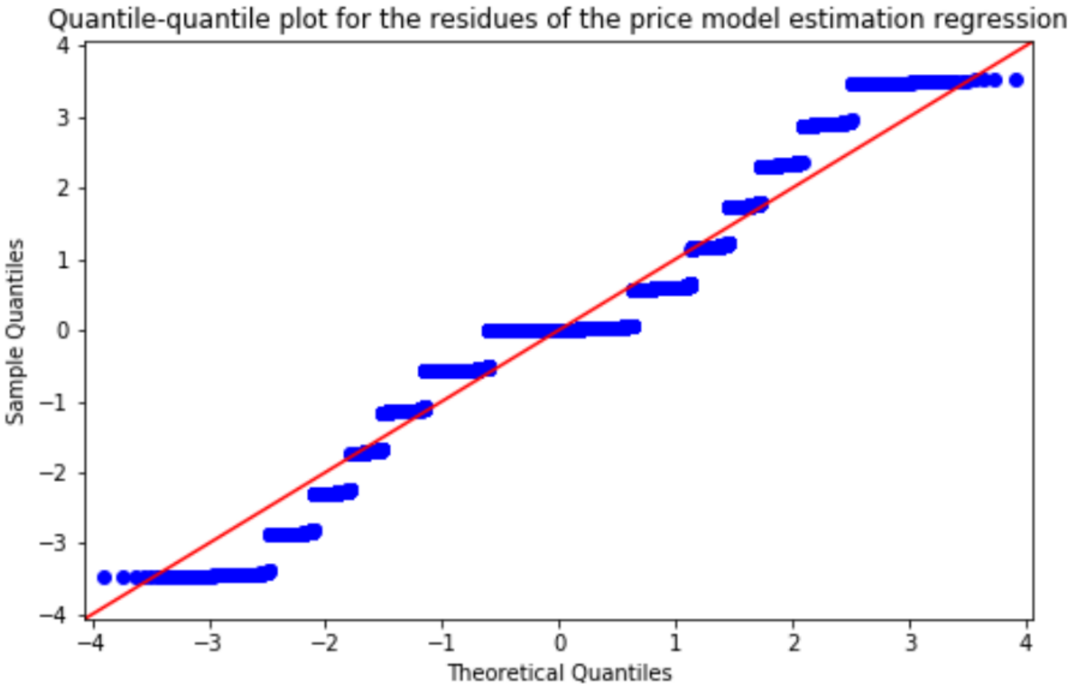

Figure 5: Quantile-quantile plot for the price model estimation regression residuals. The theoretical distribution is the standard normal distribution. Data INFY, 8th jan, 2020

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.112
Model:                            OLS   Adj. R-squared:                  0.112
Method:                 Least Squares   F-statistic:                     5741.
Date:                Wed, 08 Jul 2020   Prob (F-statistic):               0.00
Time:                        20:47:02   Log-Likelihood:            -1.4015e+05
No. Observations:               45328   AIC:                         2.803e+05
Df Residuals:                   45326   BIC:                         2.803e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1           -32.7924      0.433    -75.767      0.000     -33.641     -31.944
const         10.4079      0.027    390.205      0.000      10.356      10.460
==============================================================================
Omnibus:                    70740.063   Durbin-Watson:                   0.258
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4241.072
Skew:                          -0.413   Prob(JB):                         0.00
Kurtosis:                       1.750   Cond. No.                         17.3
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 6: Regression summary for the buy side of the order book. Data INFY 8th Jan, 2020

2.2.1 Formulation:

- **Buy side:**

$$q_b(p,t) = A_b \exp B_b(p - p_t) \quad A_b \geq \quad and \quad B_b \geq 0 \tag{11}$$

  $q_b(p,t)$ is the number of buy bids at price p at time t.

- **Sell side:**

$$q_s(p,t) = A_s \exp B_s(p - p_t) \quad A_s \geq 0 \quad and \quad B_s \leq 0 \tag{12}$$

  $q_s(p,t)$ is the number of assets on sale at price p at time t.

.

The mid price is p_t at time t. Since in the order book logs we just had last traded price and the best ask and bid price. We had to run a log linear regression with dependent variable inventory amount and independent variable price depth on both sides of the mid price to estimate the different parameters.

2.2.2 Empirical evidence:

The above regression had very significant value (p value and F statistics less than .1%) with signs of the parameters as expected. Following are the value for INFOSYS data on 8th Jan, 2020 (see figure 6 for the buy side and figure 7 for the sell side).

2.3 Inventory process:

We denote the market maker's inventory process as L_t at time t. Using 6 to be able to write the inventory process, we had to design the λ process first. To capture the insight of [9] we put the following requirements on λ_{n+1}:

$$\begin{aligned} \mathbb{E}[\lambda_{n+1}|\mathcal{F}_n] &= \rho_n(s_n) f_n(s_n) \\ \mathbb{E}[\lambda_{n+1}^2|\mathcal{F}_n] &= (f_n(s_n))^2 \end{aligned} \tag{13}$$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.112
Model:                            OLS   Adj. R-squared:                  0.112
Method:                 Least Squares   F-statistic:                     5741.
Date:                Wed, 08 Jul 2020   Prob (F-statistic):               0.00
Time:                        20:46:24   Log-Likelihood:            -1.4015e+05
No. Observations:               45328   AIC:                         2.803e+05
Df Residuals:                   45326   BIC:                         2.803e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1           -32.7924      0.433    -75.767      0.000     -33.641     -31.944
const         10.4079      0.027    390.205      0.000      10.356      10.460
==============================================================================
Omnibus:                    70740.063   Durbin-Watson:                   0.258
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4241.072
Skew:                          -0.413   Prob(JB):                         0.00
Kurtosis:                       1.750   Cond. No.                         17.3
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 7: Regression summary for the sell side of the order book. Data INFY 8th Jan, 2020

Here ρ_n and f_n are positive continuous processes, with $\rho_n \in [0, 1]$. This is an unpredictable form of linear price impact, in the sense that, ex-post, the price increment is a linear function of the traded volume. With the above constraints, calculating the predictable quadratic variation of L_t:

$$\sum_{k=1}^{n-1} f_k^2(s_k)\mathbb{E}[\Delta_k p^2 | \mathscr{F}_k] \tag{14}$$

and predictable quadratic co-variation of L_t and p_t is:

$$-\sum_{k=1}^{n-1} \rho_k(s_k) f_k(s_k)\mathbb{E}[\Delta_k p^2 | \mathscr{F}_k] \tag{15}$$

Note that the price process (p_t) is adapted to the filtration $\mathbb{F}_t$. As a result of the above observations, we model L_t as:

$$\begin{aligned} dL_t &= -\rho_t(s_t) f_t(s_t) d\mu_t + f_t(s_t)\sigma_t dW_t^{'} \\ dL_t &= -\rho_t(s_t) f_t(s_t) dp_t + f_t(s_t)\sigma_t \sqrt{1 - \rho_t^2(s_t)} dW_t^{\perp} \end{aligned} \tag{16}$$

Both the above models are the same. One version is in terms of the price process. The noise process $W_t^{\perp}$ is independent of W_t.

2.4 Function f (f_t):

2.4.1 Formulation:

To design the function f, we can think of it as the negative of the change in inventory per unit change in price. From the point of view of an order book, price (mid-price) changes because trades occur at some different price level than the last. Since we only have one

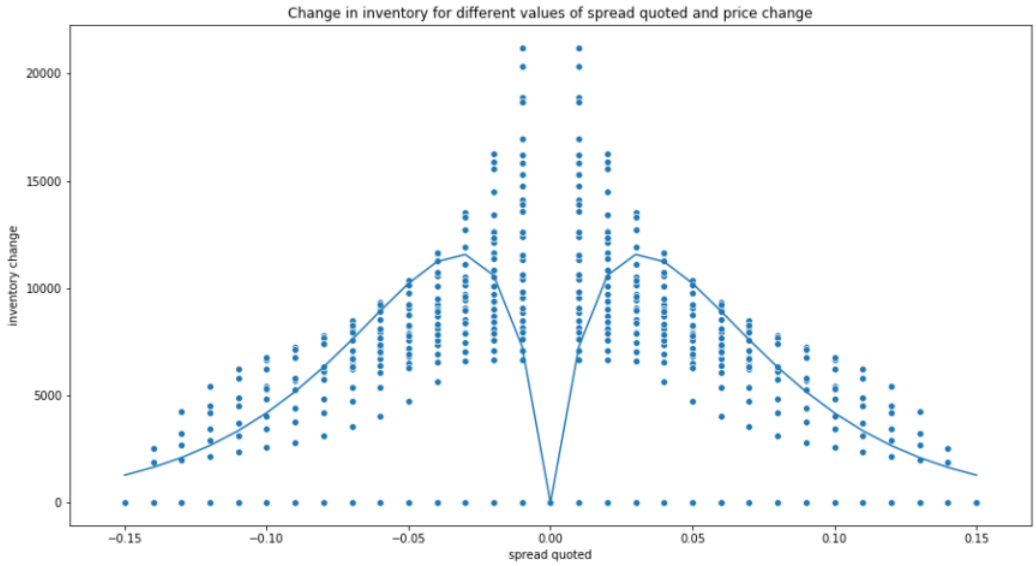

Figure 8: Estimating f_t. The blue line is the estimated function.

market-maker participating in the market, only she can make limit-orders execute. The way she does that is by the spread (s_t) she quotes, executing all the limit-orders that falls under her quotes. Since we have already assumed the order book to look the same through out time (i.e. the order book shape is constant), once we fix the spread quoted by our lone market-maker, we can determine the price level at which limit orders were executed. Because the only way limit-orders are executed is through our market-maker. About the market-orders we can have two school of thoughts to fit our theory. One can be to see market-orders as limit-orders at the best available price, or the second could be to assume market-orders on the both sides (buy and sell) roughly cancel each other's effect on the change in price level trades get executed and whats left is the noise. We any ways are going to need to account for noise. The function ρ_t takes care of the noise and it is discussed in the next section.
The function f_t is the ratio of ΔL_t and Δp_t for a quoted s_t. Now trouble is, there will be a bunch of values corresponding to different Δp_t that function f_t should take for every s_t. Since that is impossible, we do what is the next best thing, we take the best fit line. Following is the form we believe f_t should take:

$$f_t(x) = a|x|e^{-b|x|} \tag{17}$$

2.4.2 Empirical evidence:

Following is a plot illustrating the estimation of the function f(figure 8). The regression summary is figure 9.

2.5 Rho (ρ_t):

2.5.1 Formulation:

There is a source of noise to ΔL_t in addition to the one already mentioned in the previous section given that we are given Δp_t. One is from the fact that the order-book might not always look like our proposed shape, or the least there might be some deviation from the

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.903
Model:                            OLS   Adj. R-squared:                  0.902
Method:                 Least Squares   F-statistic:                     1933.
Date:                Thu, 09 Jul 2020   Prob (F-statistic):          2.83e-107
Time:                        13:30:35   Log-Likelihood:                -79.765
No. Observations:                 210   AIC:                             163.5
Df Residuals:                     208   BIC:                             170.2
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1           -31.7281      0.722    -43.969      0.000     -33.151     -30.306
const         13.8146      0.046    302.701      0.000      13.725      13.905
==============================================================================
Omnibus:                       11.028   Durbin-Watson:                   0.262
Prob(Omnibus):                  0.004   Jarque-Bera (JB):               11.243
Skew:                           0.545   Prob(JB):                      0.00362
Kurtosis:                       3.314   Cond. No.                         29.5
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 9: Regression summary for estimating f_t.

shape. As a result of which the ratio of $\frac{-\Delta L_t}{\Delta p_t}$ that we use to relate the actual ΔL_t might not come out to be precise. So we use the function ρ_t as a scaling to adjust the theoretical value to the actual. In theory the function ρ_t should be:

$$d[W, W']_t = \int_{u=0}^{u=t} \rho_u(s_u) du \tag{18}$$

For actual calculation we used the correlation of the two processes (W and W') to estimate ρ. For every time t, we took a window say $s \in [t-\delta, t+\delta]$ and calculated the correlation of the two processes W and W' in the window to estimate $\rho(s_t)$. Once we had done so for all the points $s \in [t_0, t_f]$, we used linear regression to estimate ρ. That is we used the following form for $\rho()$:

$$\rho_t(x) = c_\rho + m_\rho x \tag{19}$$

2.5.2 Empirical evidence:

See figure 10 for the final run for the estimation of ρ_t on the INFY 8th jan, 2020 data. The function comes out to be a constant function.

3 The stochastic optimal control problem:

3.1 Problem formulation:

The market maker's wealth process (X_t) can be seen as comprising the cash she is carrying and the cash she would generate by liquidating her entire inventory (L_t) at price p_t. Our objective could be to maximise our expected terminal wealth, i.e. $\mathbb{E}[X_t]$. Let's describe the evolution of the wealth process in the discrete setting. We can then take appropriate limiting case to convert it into the continuous setting.

$$\begin{aligned} \Delta X_t &= \Delta(L_t p_t) + \Delta K_t \\ \Delta K_t &= -p_t \Delta L_t + \frac{s_t}{2}|\Delta L_t| \end{aligned} \tag{20}$$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.002
Model:                            OLS   Adj. R-squared:                  0.002
Method:                 Least Squares   F-statistic:                     43.58
Date:                Sun, 12 Jul 2020   Prob (F-statistic):           4.15e-11
Time:                        01:00:27   Log-Likelihood:                 64436.
No. Observations:               22263   AIC:                        -1.289e+05
Df Residuals:                   22261   BIC:                        -1.289e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1            -0.0005   6.88e-05     -6.602      0.000      -0.001      -0.000
const          1.0001      0.000   9988.410      0.000       1.000       1.000
==============================================================================
Omnibus:                    80213.470   Durbin-Watson:                   0.252
Prob(Omnibus):                  0.000   Jarque-Bera (JB):      28480827925.412
Skew:                         -74.365   Prob(JB):                         0.00
Kurtosis:                    5542.028   Cond. No.                         1.85
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 10: Regression summary for estimating ρ_t.

The spread quoted at time t is s_t. In the continuous setting, the above equation simplifies to [5, see] the following:

$$dX_t = d(L_t p_t) - p_t dL_t + \frac{1}{\sqrt{2\pi}} \sigma_t s_t f_t(s_t) dt \tag{21}$$

The above is the state dynamics. With three state variables:

- Wealth (X_t)
- Inventory (L_t)
- Price (p_t)

Our control variable is $\mathbf{s_t}$. The state dynamics are 21, 7 and 16.
Our objective function becomes:

$$\max \mathbb{E}[X_{t_f}] \tag{22}$$

The initial conditions would be the following:

- p_0 to be some given constant.
- $L_0 = 0$
- $X_0 = 0$

3.2 Solution:

Seeing the above structure of the optimal control problem, it can be solved by using the Pontryagin's maximum principal [7, 5, see]. The resulting hamiltonian is the following:

$$\mathbb{H}_t(s, L, Y, Z, Z^{\perp}) = -\rho_t(s_t) f_t(s_t)[(Y_t - p_t)\mu_t + \sigma_t Z] + \frac{\sigma_t^2}{\sqrt{2\pi}} s_t f_t(s_t) + \sigma_t f(s_t) \sqrt{1 - \rho^2(s_t)} Z^{\perp} \tag{23}$$

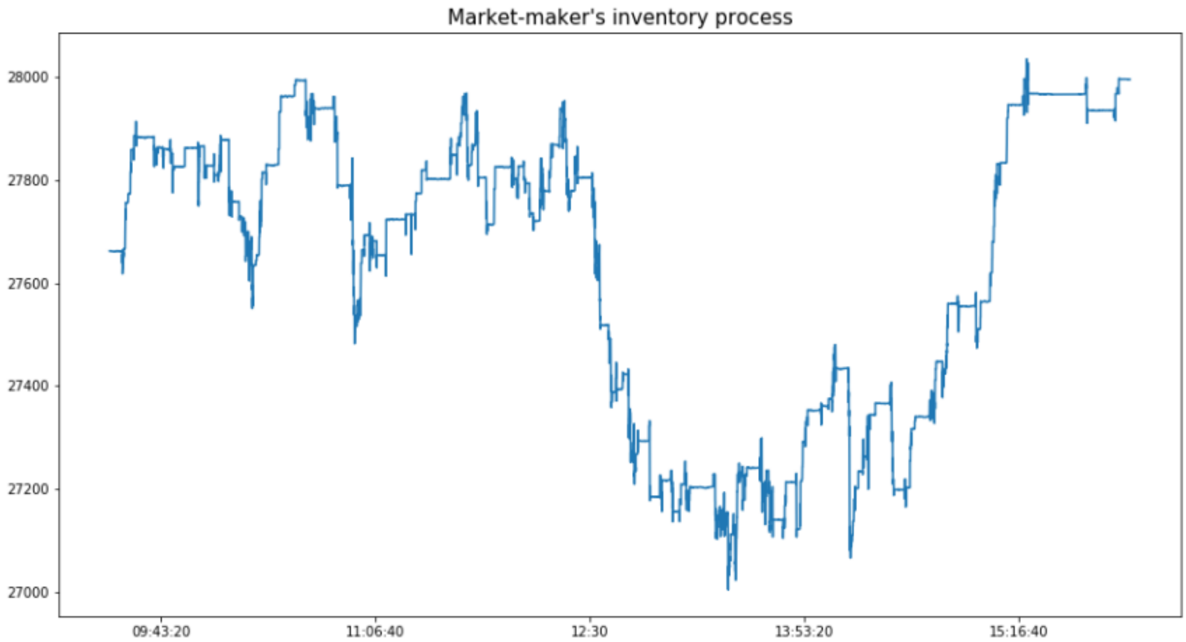

Figure 11: Market-maker's inventory process corresponding to YESBANK on 10th Jan, 2020

and the above expression is maximised by setting

- $Z_t^{\perp} = 0$
- $Y_t = \mathbb{E}[p_{t_f} | \mathcal{F}_t]$
- The optimal control is:

$$\begin{aligned} s_t &= \arg\max_{x \in [0,\infty)} \frac{x}{\sqrt{2\pi}} f(x) - \alpha_t \rho(x) f(x) \\ \alpha_t &= \mathbb{E}[p_{t_f} - p_t | \mathcal{F}_t] \frac{\mu_t}{\sigma_t^2} + \frac{Z_t}{\sigma_t} \end{aligned} \tag{24}$$

And the expected earning at the end of the trading window is:

$$\mathbb{M} = \mathbb{E}\left[\int_0^{t_f} (\max_{x \in [0,\infty)} \frac{x}{\sqrt{2\pi}} f(x) - \alpha_t \rho(x) f(x)) \sigma_t^2 dt\right] \tag{25}$$

4 Inferring the solution:

Since we observe very less variance in ρ across time as it is mostly close to 1 all the time, equation 24 can be seen to vary only in α_t across time. So α_t is the variable which decides the spread quoted. In our case, the optimum solution always happen to be the corner points. Either the market-maker quotes the maximum spread ($s_{max} = 5$ for us) or the minimum ($s_{min} \in [0, 0.3]$ for us).

We notice that whenever $\alpha_t > 0$ there's a tendency to quote the maximum spread. And whenever $\alpha_t < 0$ there's a tendency to quote the minimum spread. There are two parts to α_t :

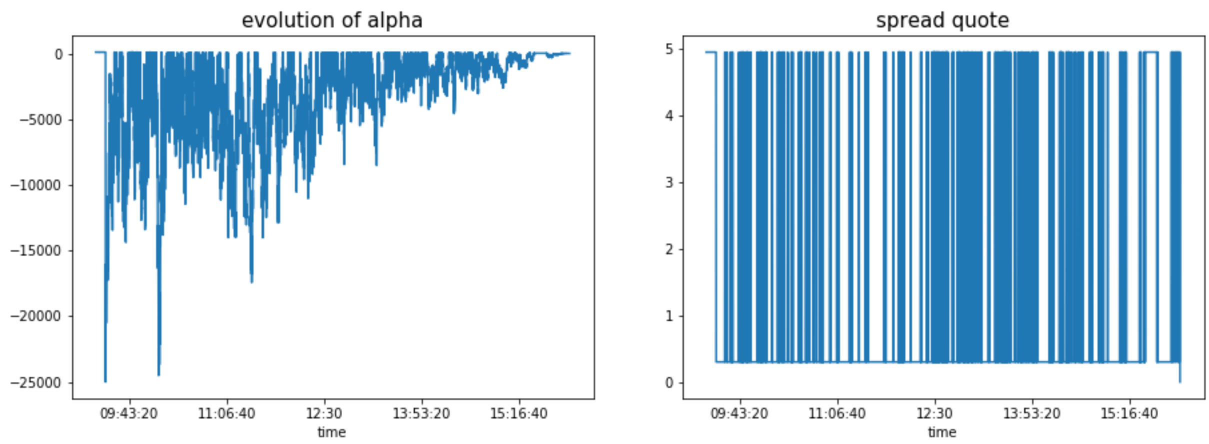

Figure 12: Final calculation of α_t and s_t for YESBANK, 19-09-2019 data.

- Product of expected net price movement and present price movement ($\mathbb{E}[p_{t_f} - p_t|\mathscr{F}_t]\frac{\mu_t}{\sigma_t^2}$)
- the volatility of the net price movement ($\frac{Z_t}{\sigma_t}$)

A small spread is quoted when the first quantity is negative enough to counter the second quantity (second quantity is always positive). The first quantity is negative when the current direction of price movement is opposite to the expected net price movement, i.e. market maker expects a price reversal in the near future. A large spread can be quoted when the volatility in the net price movement is high. See also figure 12 below. A consequence of equation 25 is that the market maker is on average short α_t and, for α_t being fixed, long volatility (σ_t).

5 Defining the algorithm:

- ***Step 1*** Estimate the order book shape.
- ***Step 2*** Using the order book shape come up with function (f_t)
- ***Step 3*** Estimate the parameters for the price model.
- ***Step 4*** Using the price model, simulate N (=1000 for us) paths.
- ***Step 5*** Using the synthetic N paths come up with $\mathbb{E}[p_{t_f} - p_t]$ and Z_t for all t. Take all the N values for each point t and get average or standard deviation corresponding to each t.
- ***Step 6*** Using the above expectation and Z_t find the optimal spread quotes every second (s_t)

- ***Step 7*** Assume some default values for the function ρ_t.

- ***Step 8*** Using defined ρ_t and f_t calculate the optimal spreads to be quoted by the market-maker.

- ***Step 9*** Corresponding to the quoted spread and Δp_t observed (we have used $\Delta =$ 1 trade period), calculate the resulting inventory changes. Using the ΔL_t, calculate the $\Delta W_t^{'}$.

- ***Step 10*** For every time point t, define $\rho_t^{'}$ as the correlation of the series ΔW and $\Delta W^{'}$ in the time period $[t-\delta, t+\delta]$. Take the values of δ to be around the range of 100 to 1000 (we took 1000) as per the computational feasibility.

- ***Step 11*** Using $\rho_t^{'}$ estimate $\rho^{'}$ using linear regression. In our case we assumed ρ to be linear in nature.

- ***Step 12*** Go back to ***Step 6*** if there is sufficient difference between ρ_t and $\rho_t^{'}$. Else continue.

- ***Step 13*** Calculate the expected profit. And call it $\mathbb{M}(\omega_1)$.

- ***Step 14*** Simulate n(=50 for us) possible realizations of the price process using the parameters estimated in ***Step-3*** $(\omega_1, \omega_2 ..., \omega_n)$, repeat ***Step-6*** to ***Step-12*** for each of the ω. Calculate $\mathbb{M}(\omega_1), \mathbb{M}(\omega_2), ...\mathbb{M}(\omega_n)$ and estimate the distribution of the expected payoff ($\mathbb{M}$).

6 Conclusion:

We proposed a price dynamics (see equation 7) during a fire sale. Our motivation was the idea that during a price crash the movement is majorly sentimental and is driven by panic. The model technically represents the following idea, the greater a price dip public sees, the further it drives down the price. We empirically tested the price model on NSE data for top 5 NIFTY performers corresponding to days when they fell substantially (greater than 2%). For the study we used a total of 10 events (company - day combination).
In our model, since we have just one market-maker, we could model our market-maker's change in inventory as a linear function of the price change. It follows from the idea that any limit-order from any market participant has to be executed by the market-maker only. And any price change is a result of the execution of market-orders or best placed limit-orders (which again can be seen as market-orders), because otherwise the orders would remain un-executed and have no affect on the "last traded price".
Once we had a model for evolution of all the concerned state variable, we went on to define our objective function which is to maximise our expected terminal wealth. This part of our work followed mechanically as per the framework proposed by [5]. Once the control problem was modeled and solved, we went on to test it on the aforementioned (in Table below) events.

Table: Events analysed - Company-Day combinations

Company	date	maximum-fall
YES BANK	19/09/2019	-17.19%
YES BANK	10/01/2020	-8.30%
WIPRO	26/11/2019	-3.28%
WIPRO	01/08/2019	-2.96%
TECH MAHINDRA	09/01/2020	-1.91%
TECH MAHINDRA	06/11/2019	-2.17%
INFY	06/01/2020	-2.66%
INFY	08/01/2020	-2.07%
AXIS BANK	08/01/2020	-5.05%
AXIS BANK	07/11/2019	-3.17%

Since the problem at hand is stochastic in nature, the price process we see is just one out of infinite realisations (ω_i). Once we had estimated the model, we simulated more synthetic ω. And applied our algorithm on all the synthetic paths the price process could have taken. The resulting distribution of the terminal wealth is inferred as the distribution for expected payoff upon application of our algorithm. The spread of the distribution gives an idea of the risk involved with the algorithm. Empirically, the distribution appears normal with mean and variance specific to the events. See figure 13.
Future work can be in terms of improving the price model. A more insightful model would allow the optimization model a more accurate decision making capability. The proposed framework can be applied to other markets and environment as well. Challenges include modeling the interaction of different components correctly. Most fundamental component is the order book shape, which once determined correctly gives direction to the modeling of the different components.

References

[1] Robert Almgren and Neil Chriss. "Optimal Execution of Portfolio Transactions". In: *Journal of Risk* 3.2 (2000), pp. 5–39.

[2] Sebastian Jaimungal Álvaro Cartea and José Penalva. *Algorithmic and High-Frequency Trading*. Cambridge University Press, 2015.

[3] Marco Avellaneda and Sasha Stoikov. "High-frequency trading in a limit order book". In: *Quantitative Finance* 8.3 (2007), pp. 217–224.

[4] Rene Carmona and Kevin Webster. "The self-financing equation in high frequency markets". In: *arXiv preprint arXiv:1312.2302* (2013).

[5] René Carmona and Kevin Webster. "The self-financing equation in limit order book markets". In: *Finance and stochastics* 8.3 (2019), pp. 729–759.

[6] Larry Eisenberg and Thomas H. Noe. "Systemic Risk in Financial Systems". In: *Journal of Risk* 47.2 (2001), pp. 236–249.

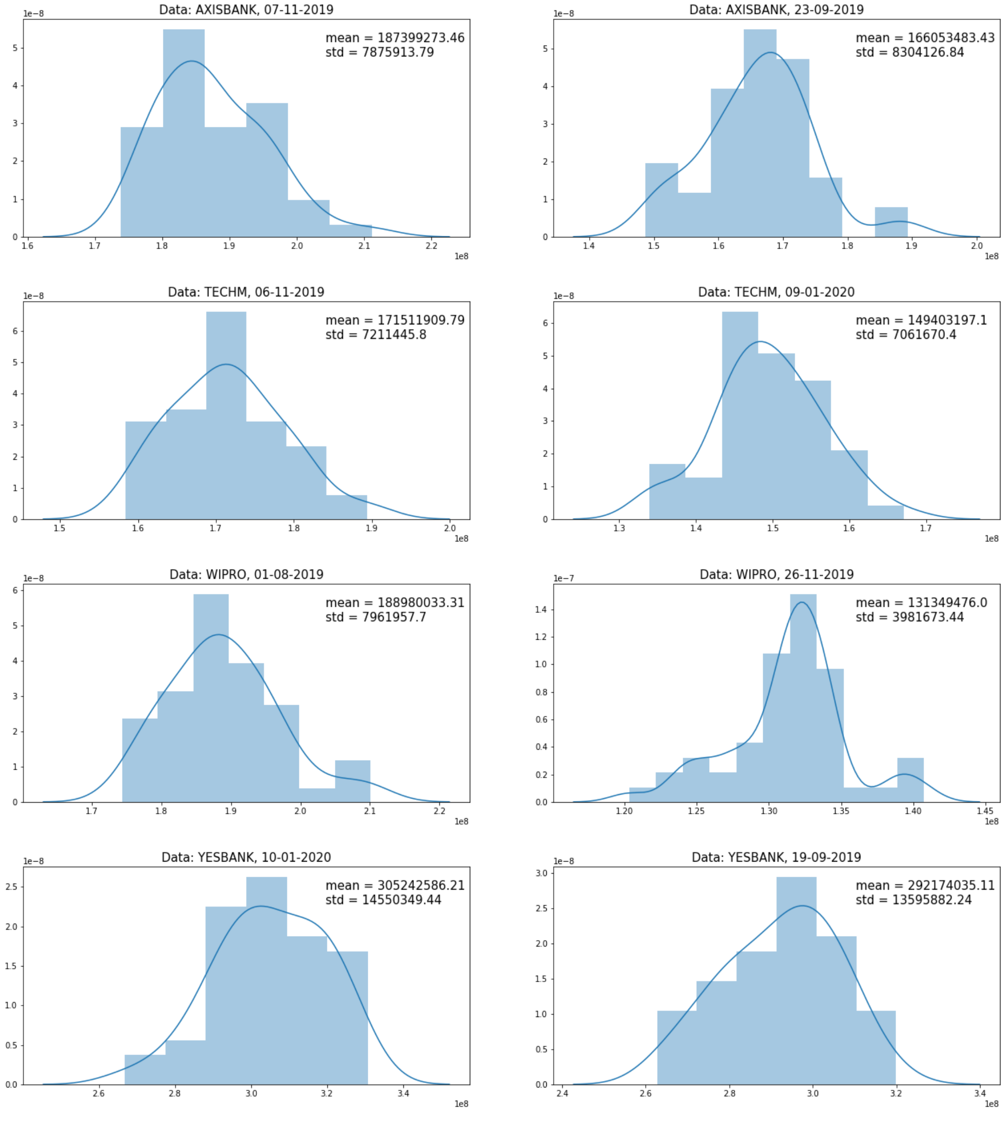

Figure 13: Distribution of expected payoffs for 50 possibilities of the path process corresponding to 10 company-day calibration from NSE data

[7] Harold J. Kushner and Fred C. Schweppe. "A maximum principle for stochastic control systems". In: *Journal of Mathematical Analysis and Applications* 8.1 (1964), pp. 287–302.

[8] Steven E. Shreve. *Stochastic calculus for finance II*. 2004.

[9] Sasha Stoikov and Mehmet Sağlam. "Option market making under inventory risk". In: *Review of Derivatives Research* 12.1 (2009), pp. 55–79.

[10] Jiongmin Yong and Xun Yu Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. 1999.