APPLIED FINANCE LETTERS

VOLUME 05, ISSUE 02, 2016

MULTIDIMENSIONAL LIQUIDITY: EVIDENCE FROM THE INDIAN STOCK MARKET

SHARAD NATH BHATTACHARYA[1*], PRAMIT SENGUPTA[2], MOUSUMI BHATTACHARYA[3], BASAV ROYCHOUDHURY[4]

1. Indian Institute of Management Shillong, Meghalaya, 793014 India
2. Army Institute of Management Kolkata, West Bengal, 700027 India
3. Indian Institute of Management Shillong, Meghalaya, 793014 India
4. Indian Institute of Management Shillong, Meghalaya, 793014 India

* Corresponding Author: Sharad Nath Bhattacharya, Indian Institute of Management Shillong, Nongthymmai, Shillong, Meghalaya, 793014 India. ✉ snb@iimshillong.ac.in

Abstract: Various dimensions of liquidity including breadth, depth, resiliency, tightness, immediacy are examined using BSE 500 and NIFTY 500 indices from Indian Equity market. Liquidity dynamics of the stock markets are examined using trading volume, trading probability, spread, Market Efficiency coefficient, and turnover rate as they gauge different dimensions of market liquidity. We provide evidences on the order of importance of these liquidity measures in the Indian stock market using machine learning tools like Artificial Neural Network (ANN) and Random Forest (RF). Findings reveal that liquidity variables collectively explain the movements of stock markets. Both these machine learning tools perform satisfactorily in terms of mean absolute percentage error. We also find a lower level of liquidity in the Bombay Stock Exchange (BSE) than the National Stock Exchange (NSE) and findings supports the liquidity enhancement program recently initiated by BSE.

Keywords: Liquidity, Turnover Rate, Market Efficiency Coefficient, Trading Probability, Artificial Neural Network, Random Forest.

1. Introduction

Liquidity is often explained as the ability to do large transactions, quickly, at low transactions costs and the evidences on the relation between liquidity and returns is important due to the fact that if liquidity affects returns, then from an investor's point of view liquidity risk needs to be priced. The most influential work on this front owes to Amihud and Mendelson (1986), who provide the first theoretical motivation establishing the relation between assets with low liquidity (or high transaction costs) and return premium. Their model was single-period with non-stochastic levels of liquidity. However, in multi-period models (Constantinides, 1986; Heaton & Lucas, 1996), it has been shown that cross-sectional differences in liquidity are not a pre-condition for a large premium on liquidity. There has been a resurgence of interest in the time-series dynamics of liquidity as well as the impact of the level of liquidity and liquidity risk on expected returns and, in turn, the cost of capital. An important observation about liquidity is that it is a parameter often endogenous to the environment. The interaction between investors' buying and selling decisions determines liquidity in equilibrium. Given the endogeneity of liquidity, it is of particular interest to explore the nexus between financial market movements and time-series movements in liquidity.

Liquidity is one of the imperative characteristics of a financial market and is considerably important for investment plans and financial assets. It probably does not have a single universally accepted definition. It changes with asset class and type of markets. Even within various financial markets, liquidity is empirically characterized in terms of breadth, depth, and resilience, often along with tightness and immediacy. The liquidity of major world financial markets substantially varies over time. Thus the unpredictability of market liquidity thereby is an important source of risk for investors.

In 1996 the NSE was set up, but other institutions and regulations facilitating trade like clearing corporations, depository and dematerialization, elimination of badla - a charge, which the investor pays for carrying forward his position, rolling settlement, ETF and derivatives trading through NSE were set up subsequently. Additionally, post 2000 we experienced events like the IT boom, stock market scams and World recession due to global financial crises. It would be interesting to see how liquidity has changed over time after these developments and events. The goal of this paper is to explore whether the Indian Stock market is related to its endogenous liquidity measures. We test for liquidity in terms of market depth, breadth, and resilience by using different liquidity measures that are deemed appropriate for equity market.

The next section provides some details on machine learning tools used in this study and in section 3 we present the previous research on similar and allied topics. In section 4 and 5 we discuss the data and methodology. In section 6, we discuss the findings and analysis and in section 7 we present our conclusion.

2. Machine learning Tools: Artificial Neural Networks and Random Forest

During last few years there has been much advancement in the application of machine learning algorithms in stock market index forecasting, endeavouring extraction of patterns in the market. Patel et al. (2015) and Wu and Lee (2015) provide a good summary of the work done in this field. Their work highlights the limitations of traditional statistical models including moving average, exponential smoothing, and ARIMA models which are linear in their predictions of the future values. From a statistical point of view, Artificial Neural Networks (ANNs) are analogous to nonparametric, nonlinear, regression models. However, the traditional statistical models have limitations in understanding the relationship between the input and the output of the system, especially when the system shows chaotic behaviour and is complex. Another machine learning algorithm which has been found to be good at such predictions is Random Forest. Theofilatos et al. (2012) apply five learning classification techniques (K-Nearest Neighbours algorithm, Naïve Bayesian Classifier, Artificial Neural Networks, Support Vector Machines and Random Forests) and observe that techniques like Support Vector Machines and Random Forests clearly outperform all other strategies in terms of annualized return and Sharpe ratio. Qin et al. (2013) applied the Random Forest method (Gradient Boosted Random Forest) as a nonlinear trading model to the stock market return of Singapore stock exchange and suggested that the proposed trading methods outperformed buy and hold strategy for similar period.

2.1 Artificial Neural Network (ANN)

ANNs are data driven models which can be used for non-linear natural real world systems while linear models generally fail to understand the data pattern and analyze when the underlying system is a nonlinear one. While some parametric nonlinear model such as ARCH and GARCH models have been in use for financial forecasting, most of such nonlinear statistical techniques require that the nonlinear model be specified before the estimation of the parameters is done. This requirement limits such models, as it generally happens that pre-specified nonlinear models may fail to observe the critical features of the complex system under study. ANNs are able to independently learn the relations inherent in the input data and discover nonlinear relations in the input data set without a priori assumptions about the relation between the input and the output.

ANN is a massively parallel distributed processor made up of a simple processing unit which has a natural propensity for storing experiential knowledge and making it available for use (Haykin, 1999). They are composed of one or more hidden layers sandwiched between the input and the output layers. Each layers is made up of a given number of nodes, and in case of a simple Feed Forward Multi-Layer Perceptron (MLP) ANN, each node in a given layer is connected to the ones in the next layer by arcs knows as synapses, taking cue from biological neurons in our bodies which are connected to each other and accept electrical charges across synapses. The input layer will have as many nodes as predictor variables (which takes in the input values to the network), and the output layer will have one node for estimation models (providing the output value) or for binary classification models (providing the probability for one of the output classes). In case of multiple (more than two) output classes, the output layer will have one node for each possible output class. The hidden layers can have any given number of layers with any given number of nodes in each of them. An illustrative ANN with two hidden layers of 3 nodes each, four input nodes and one output node is shown in Figure 1.

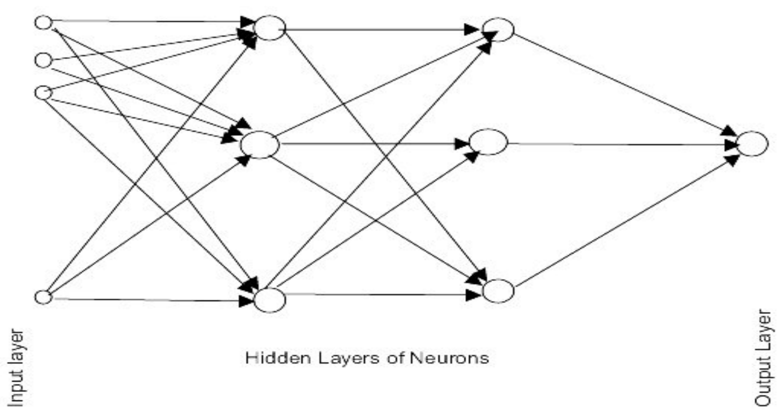

Figure 1: ANN with two hidden layers

Each arc in the network is assigned certain weight w. As an arc connects node i and node j, the value from node i gets multiplied with the corresponding weigh of the arc while traversing the concerned arc. Each node j (except those in the input layer of the ANN) also has some constant bias θj, which gets added up with the inputs received at node j; the output of node j being a function of these:

$$g(s) = g(\theta_j + \sum_{i=1}^{p} w_{i,j} x_i). \quad (1)$$

The function g(s), known as the activation function, can be a linear, exponential, or a sigmoidal function. It can be the same function at each node in the network, or there can be different activation function, say sigmoid, at the hidden nodes, and say linear, at the output node(s). During the training phase of the ANN, the network is trained in terms of deciding on the weights wi,j and the biases θj, for every i, j, so that the network can provide the desired output.

One of the learning techniques used in MLP ANNs is the backpropagation of errors. The backpropagation algorithm falls into the general category of gradient descent algorithms, which intend to find the minima/maxima of a function by iteratively moving in the direction of the negative of the slope of the function to be minimized/maximized. The main objective is to minimize the error function. In this algorithm, the weights and biases are updated on a pattern-by-pattern basis until one complete epoch has been dealt with. The adjustments to the weights are made in accordance with the respective errors computed for each pattern presented to the network. The arithmetic average of these individual weights over the entire training set is an estimate of the true change that would result from the modification of the weights based on the error function. A gradient descent strategy is adopted to minimize the error.

Thus, ANNs have a built-in capability to adapt the network parameters to the changes in the studied system. A neural network trained to a particular input data set corresponding to a particular environment; can be easily retrained to a new environment to predict at the same level of environment. However, while blessed with good predictive performance, ANN is a black box algorithm, and hence does not provide any information regarding the relative importance of predictor variables used in the model.

2.2 Random Forest

Random Forest is an ensemble method, whereby a combination of Classification and Regression Trees (CARTs) are used; with the individual outputs from each of the CARTs finally combined to generate the output for the Random Forest. The results are combined by a method of voting for classification, and by a method of averaging the individual outputs in case of regression to arrive at the final result, the latter being the one used in our models.

Each of the CARTs in Random Forest are grown randomly from the training dataset provided to train the Forest. The individual trees are grown using different training sets. A random vector Θ_k is generated to grow a tree from the training set provided to train the Random Forest. Θ_k is independent of past random vectors Θ_1, Θ_2, …, Θ_(k-1), but follow the same distribution. The training sets used to develop the various trees are derived by randomly drawing the records, with replacement, using the random vector Θ_k, from the training set originally provided for the Random Forest. A new tree is grown with each of these new training sets using random feature selection. These trees are allowed to grow without pruning. Each individual tree is thus a classifier or regressor of the form $\{h(x, \Theta_k), k = 1,2, ..\}$.

It has been shown that for a large number of trees, because of the law of large numbers, Random Forest does not overfit, instead, it produces a limiting value of generalized error. Random Forest also does not provide much insight into the model building, it does compute and provide the relative importance of predictor variables in the model.

3. Literature Review

The increasing empirical evidences on the liquidity and stock market nexus in quite voluminous. Kumar and Misra (2015) provide an excellent review of the frameworks currently available for modelling liquidity. Here we attempt to review the most influential studies in this area. Chordia et al. (2001), use trading activity and turnover rate to conclude that liquidity has a negative effect on risk-adjusted stock returns, which was supported by Pastor and Stambaugh (2003); Marshall and Young (2003) and Moore and Sadka (2006) for different markets. On the Spanish stock market, Martinez et al. (2005) observe a significant and positive relationship between the Amihud (2002) illiquidity measure and returns in both the unconditional and conditional asset pricing models. Moreover, using the Pastor and Stambaugh (2003) liquidity measure, they find a significant negative relationship in only the conditional asset pricing model. However, when they use the bid-ask spread as a proxy for liquidity, they do not find any relationship. Faff et al. (2010) report a negative association between expected stock returns and liquidity measures but contrary to perceived notion that liquidity is more important during bear phases, they observe that liquidity is priced during expansionary phase of business cycle but not significantly priced during contraction phase. This apparent consensus of a negative relation between stock-level liquidity and expected returns, a persistent negative shock to a security's liquidity should, as pointed out by Acharya and Pedersen (2005), result in low contemporaneous returns and high future returns, and vice versa, has been challenged on numerous occasions. It is argued that this prediction of a negative relation between liquidity shocks and future returns may not hold in a market in which information is not fully reflected into prices due to market frictions. Bali et al. (2014) provide evidence that stock markets underreact to stock-level liquidity shocks and liquidity shocks are not only positively associated with contemporaneous returns, but they also predict future return continuations for up to six months. Batten and Vo (2014) observe a positive relation between liquidity and stock returns for emerging equity markets which contradicts the negative correlation typically found in stock returns in developed markets obtained earlier. Most of the work on liquidity has used standard econometric techniques. However, machine learning algorithms were used by some authors for stock market prediction. While Dutta et al. (2006) evidence that ANN performs satisfactorily in predicting closing prices of SENSEX, the leading index of Bombay Stock Exchange, Qin et al. (2012) evidence support for Random Forest based trading model for the Singapore exchange. Sala (2011) develops an alternative approach of liquidity risk modelling using a recurrent neural network and shows that machine learning may be an important alternative while modelling liquidity risk. In the Indian context, Krishnan and Mishra (2013) explore liquidity patterns in the Indian stock market while Kumar and Mishra (2015) explore patterns for individual stocks, we did not find any study in Indian context that uses liquidity measures to explain stock market movements.

Clearly evidences on effects of liquidity on stock market do not seem to converge but still there is a general consensus that liquidity reduces returns, and often empirical evidence supports the idea that risks emanating from liquidity need to be priced. It follows that an investigation on whether or not liquidity risk needs to be priced on the Indian stock market offers a fresh perspective on the liquidity-return nexus and worth a review. Given the idiosyncrasies of Indian equity market, the study attempts to explore whether stock market return variations can be explained by collection of liquidity measures used in the literature and if the two major Indian stock exchanges NSE and BSE differs in terms of liquidity. Also there is a natural need to vouch and verify the existing research findings especially with emergence of changing microstructure.

4. Data

In this paper, to gauge the robustness of the effect of liquidity on returns, we consider five liquidity measures. Following Korajczyk and Sadka (2008), we use trading volume and the turnover rate as measures of liquidity. Amihud and Mendelson (1986) suggested the strong theoretical background for the use of the turnover rate arguing that liquidity is correlated with trading frequency in equilibrium, and is well discussed in Datar et al. (1998). The turnover considered here is the ratio of monthly trading volume and market capitalization. In addition, we follow Narayan and Zheng (2011) and consider the trading probability as an additional measure of liquidity, which is calculated as:

$$\textit{Trading Probability (Tp)} = 1/\,(1 + \textit{the number of non-trading days in a month}) \qquad (2)$$

They used this measure to capture the speed dimension of liquidity and avoid the bias effects from the noise in the market as a noisy market have more risks of serial correlation effects.

We also consider the spread (high minus low) that captures the transaction costs and market efficiency coefficient (MEC) for resiliency. MEC measures the impact of execution costs on price volatility over short horizons and compares the long-term variance with the short-term variance. The variance of transaction prices are expected to be smaller in a liquid market. MEC is calculated as:

$$\text{MEC} = \frac{\text{Long Term Variance}}{\text{T} \times \text{Short Term Variance}} \qquad (3)$$

where T is the number of sub periods into which longer periods of time can be divided. We considered 5 days as short period and 30 days as long period i.e., T = 6. When MEC is less than 1 but close to it, it suggests that the market is resilient and minimum price volatility is expected.

The study focuses on two major stock exchanges of India – National stock Exchange (NSE) and Bombay stock Exchange (BSE) and considers two composite indices NIFTY 500 and BSE 500. The indices values are taken into their natural logarithm form (lnindex). The idea is to consider a well-diversified index from each exchange and so that it consists of companies of different market capitalization and categories (types).

5. Methodology

The time period considered is July 2002 to February 2016. Time series data are obtained from Bloomberg and liquidity variables are calculated. We first study the descriptive statistics of all the variables considered including their time series characteristics. Then we used machine learning techniques - Artificial Neural Network and Random Forest - to explore as to whether stock market is related to the liquidity measures considered. ANN and RF are arguably the most frequently used machine learning algorithm, and can learn any linear or non-linear function. Given the dynamic nature of the system under study, machine learning suits better than other traditional models in predicting the stock market as it can change its network parameters (synaptic weights and node biases) in real time. A feed forward neural network with standard backpropagation leaning function was used to learn the following:

$$\text{lnIndex}=f(\text{tp},\text{lnsp},\text{mec},\text{lntv},\text{Turnover Rate}) \quad (4)$$

Also, since RF has been found to have good predictive power in case of non-linear data and can learn the relationship from the data without any a priori knowledge of such relationships as in case of ANN, RF was used to learn the relationship in the above equation. In case of the machine learning algorithms of ANN and RF, the available monthly records for NSE and for BSE were partitioned into two partitions each – one for training the ANN and RF models, and the other for evaluating the performance of the trained model using the remaining data. The training partition was built by randomly picking up 70% of the records, without replacement, from the available data. All the input and output variables are contemporaneous aiming to explore a possible relationship between the liquidity dimensions and stock market movements. The training partition for NSE data contained the records of the same data as those in the training partition for BSE data. The same was the case with the partitions created for validating the models for BSE and NSE data, respectively.

Further, we tested whether liquidity in BSE and NSE are different in terms of the parameters used in this study. Hence, we tried non-parametric tests under the null hypothesis that two independent samples are from populations with the same distribution by using the Wilcoxon rank-sum test and the Kolmogorov-Smirnov test for equality of distribution functions to explore whether there is some level of equality in terms of liquidity parameters.

6. Findings & Analysis

6.1 Descriptive Statistics

Table 1: Descriptive Statistics of the variables (NIFTY 500)

	lnindex	Lnsp	Lntv	Mec	Tp	Turnover Rate
Mean	8.05	5.60	22.66	0.57	0.09	0.45
Median	8.28	5.73	22.97	0.35	0.09	0.42
Maximum	8.89	7.51	23.87	3.42	0.12	1.06
Minimum	6.54	3.36	20.59	0.09	0.07	0.21
Std. Dev.	0.62	0.72	0.79	0.53	0.01	0.13
Skewness	-0.94	-0.58	-0.94	2.21	0.14	1.28
Kurtosis	3.02	3.36	2.89	8.96	2.53	4.95
Jarque-Bera	23.91***	10.38***	24.52***	375.91***	2.08	71.02***

Note: The Table 1 above shows the mean, median, range, standard deviation and the third and fourth moments of the independent and dependent variables related to NIFTY 500. Except for trading probability, the Jarque-Bera statistics are significant for all series at 1% level (denoted by ***) indicating rejection of null hypotheses of normal distribution for these series.

Table 2: Descriptive Statistics of the variables (BSE 500)

	Lnindex	Lnsp	Intv	Mec	Tp	Turnover Rate
Mean	8.50	6.07	27.22	0.66	0.09	0.31
Median	8.75	6.21	26.98	0.45	0.09	0.21
Maximum	9.35	7.96	29.42	5.39	0.13	1.11
Minimum	6.96	3.81	25.74	0.09	0.07	0.04
Std. Dev.	0.63	0.73	0.80	0.69	0.01	0.03
Skewness	-0.94	-0.61	0.67	3.67	0.10	0.88
Kurtosis	2.99	3.35	2.82	21.43	2.36	2.56
Jarque-Bera	24.51	11.27	12.49	2691.61	3.12	22.74
Probability	0.00	0.00	0.00	0	0.21	0.00

Note: The Table 2 above shows the mean, median, range, standard deviation and the third and fourth moments of the independent and dependent variables related to BSE 500. Except for trading probability, the Jarque-Bera statistics are significant for all series at 1% level (denoted by ***) indicating rejection of null hypotheses of normal distribution for these series.

6.2 Measures of Liquidity (Trend Analysis)

Sarr and Lybek (2002) are in favour of using market indices as a proxy for stock market with the caveat that they cover only the important stocks. Figure 2 and 3 below shows general liquidity measures of Indian equity markets using BSE500 and NSE500 indices.

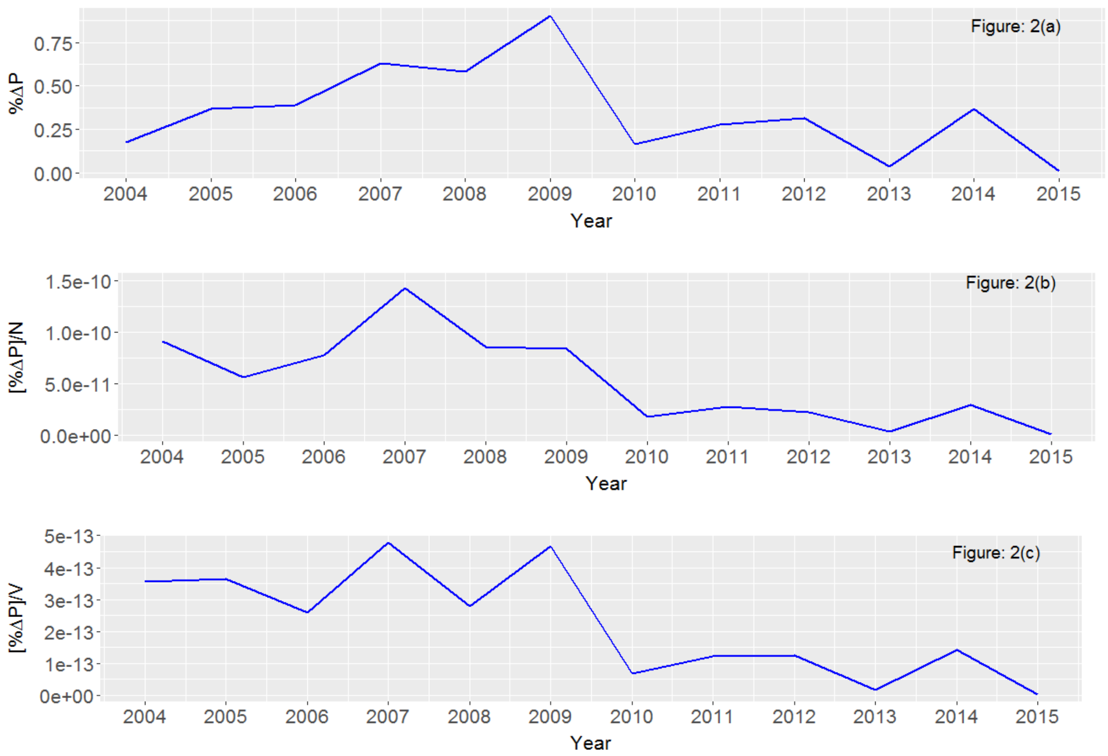

Figure 2: General Liquidity Measures – BSE 500

It is seen that in case of BSE 500 data, the volatility of the index as measured by the percent change (figure 2(a)) has increased from 2004 only to shoot up during 2007 to 2009, where equity markets around the globe were affected due to series of economic news and events post US led financial crises. The market remained flat for the majority of the time after 2010 and showed signs of volatility when India had its general election which brought a stable government in power. From 2015 onwards the market remained flat due to lack of positive global news with domestic good news being possibly nullified by negative sentiments about the Chinese economy.

The conventional liquidity ratio (figure 2(b)) relating to price changes to number of units traded have shown a upward trend since 2005, reaching its peak during 2007 and then came down till 2010 from where it fell to its lowest in 2013 and 2015.

Another conventional liquidity ratio (figure 2(c)) relating to value of transactions had its peak during 2007 only to fall in 2008, climb up again in 2009 and then sharply came down in 2010 from where it fell to its lowest in 2013 and 2015. Thus the conventional liquidity measures showed a similar type of trend from 2004 to 2007, when volatility of the index was increasing. This consistency in the behavior of conventional liquidity measures in the face of increased (decreased) price volatility can be interpreted as increase (decrease) in market depth.

After 2009, when fluctuations in the volatility of the index was observed, the conventional liquidity ratios also increased. This might be because of the reason that number of units traded (N) and turnover (V) have not experienced the same increase as before. So the possibility of a reduction in market depth cannot be ruled out post 2009.

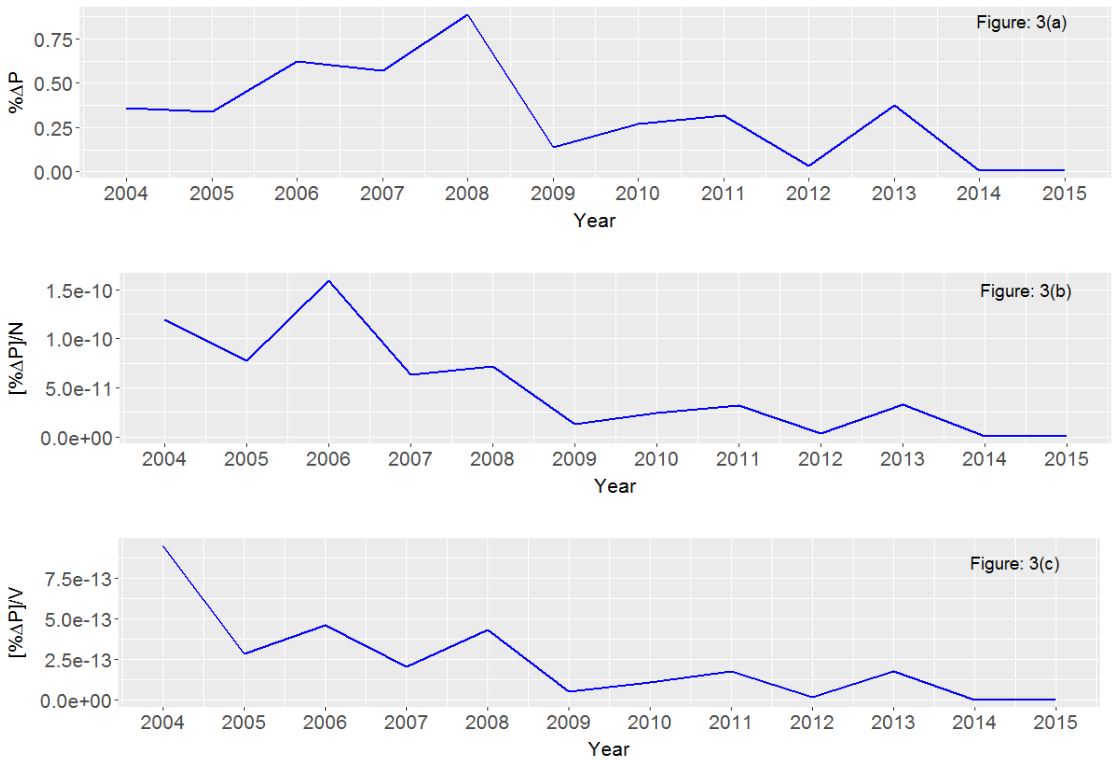

Figure 3: General Liquidity Measures – NSE 500

It is seen that in case of NSE 500 data, the volatility of the index as measured by the percent change (figure 3(a)) increased from 2004 to 2007-08, but a sharp fall was observed in 2008-09, thereafter it followed the same behavior as shown in BSE 500 series volatility.

The conventional liquidity ratio (figure 3(b)) relating to price changes to number of units traded have shown a downward trend since 2006, with sharp fall in 2007 and 2009 and was significantly low during 2014 and 2015.

Another conventional liquidity ratio (figure 3(c)) relating to the value of transactions has shown a downward trend since 2004 with sharp fall in 2007-08, and remained flat during 2014-15.

Thus the conventional liquidity measures showed a similar type of trend from 2004 to 2006, when the volatility of the index was increasing. This consistency in the behaviour of conventional liquidity measures in the face of increased price volatility can be interpreted as an increase in market depth. After 2007, when fluctuations in the volatility of the index was observed with cyclical ups and downs, the conventional liquidity ratios also started increasing. The number of shares traded (N) and turnover (V) have not experienced the same increase as before and thus the indication of a reduction in market depth during post financial crises period.

However these observations needs to be supplemented with other liquidity measures as discussed under methodology section and is reported below.

6.3 Market Efficiency Coefficient (MEC)

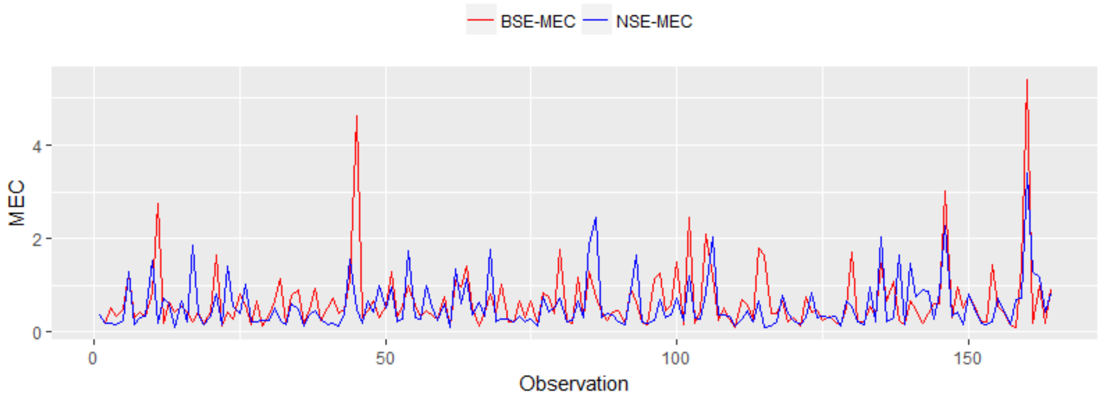

Figure 4: Market Efficiency Coefficient

The MEC exploits the fact that price movements are more continuous in liquid markets, even if new information is affecting equilibrium prices. The ratio should be closer but slightly below one in a more resilient market. MECs (BSE) are mostly around one fluctuating above and below it during the time period of the study with some outliers. So we can infer that the market was mostly resilient and a short term volatility is an expected fact when MEC is substantially below one.

MEC (NSE) was closer to one, fluctuating both above and below it with few outliers. A MEC greater than one may not be surprising as market maker intervention, inaccurate price determination involving partial adjustment to news causes prices to adjust in relatively small and positively correlated increments and this would dampen short price volatility to longer period volatility and may cause the MEC to be above one.

6.4 Turnover Rate or Turnover Velocity

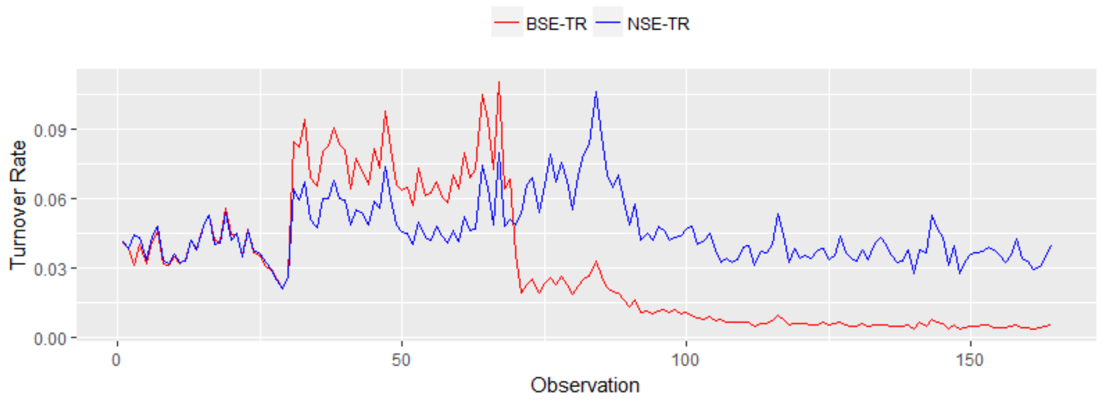

Figure 5: Turnover Rate

The Turnover rate or turnover velocity shows slightly different behaviour for BSE and NSE. At BSE, the rate was on higher side during mid-2004 till 2007 and then decreased sharply during 2008, possibly an effect of global financial crises. Since then it had a downward trend and remained low with lower fluctuations suggesting evidence of reduced breadth. In case of NSE, it picked up from mid-2004, was on the higher side till 2007and came down in 2007 and then again picked up from 2008 only to come down at around pre 2004 level and stabilized there. But it's important to note that this stabilized rate is much higher in NSE than in BBE.

6.5 Spread

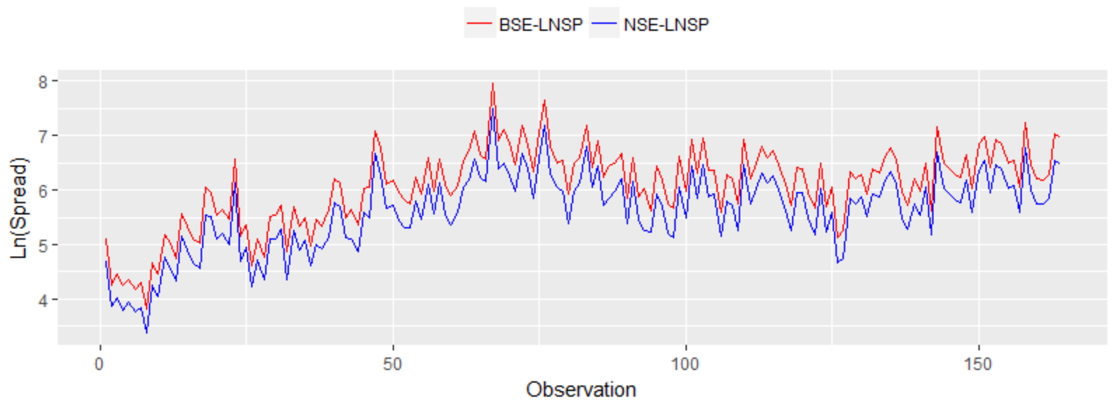

Figure 6: Spread

Trend Analysis of spread gives almost similar outcome for both the indices with spread at high levels during 2007-08 then gradually coming down with fluctuations. High spread during crises and/or world recession period led to reduced liquidity as indicated by high spread. However both BSE and NSE shows upward trends at decreasing rate.

6.6 Trading Volume

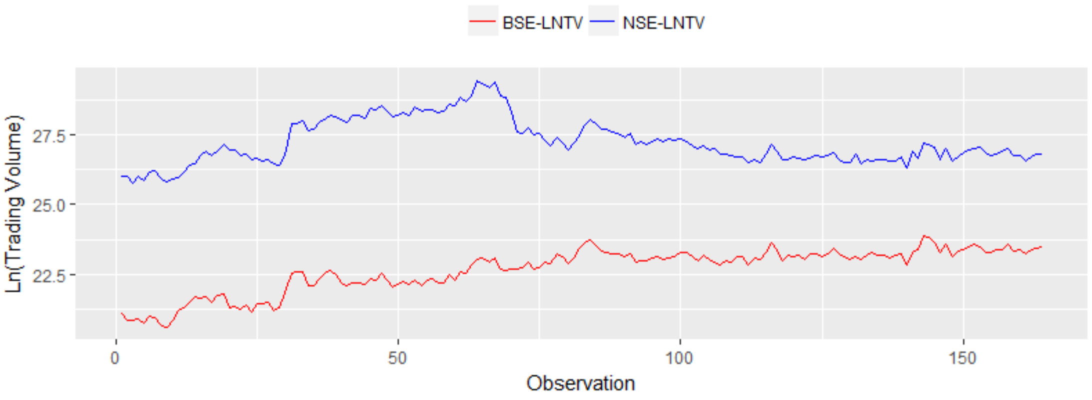

Figure 7: Trading Volume

Trading volume is a traditional measure of liquidity as Market liquidity refers to the extent to which a market allows assets to be bought and sold at stable prices. The trend analysis of trading volume of BSE and NSE gives clear indications that BSE is steadily decreasing its trading volume and liquidity while NSE's trading volume and liquidity is growing. There were obvious ups and down during global events like during financial crises, both the markets crashed but NSE picked up subsequently while the BSE could not. Even a fall in indices due to the Chinese equity meltdown and rupee crashing against the dollar in 2015 led to a spurt in trading volume as panicked investors hit the exit button.

6.7 Trading Probability

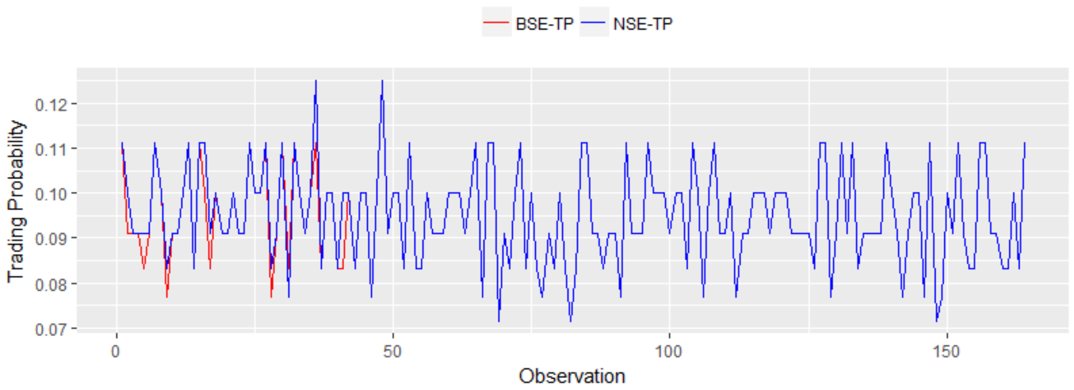

Figure 8: Trading Probability

The Trading Probability measure seems to function as an alternative to the usual logarithm of Size variable. It is expected to capture one of the dimensions of liquidity viz., Trading Speed. We observe that trading probability trend in BSE and NSE are almost similar during study period.

6.8 Artificial Neural Network (ANN)

The network used was a simple one which had one hidden layer with three nodes (H1, H2, and H3), and is represented below (Figure 9):

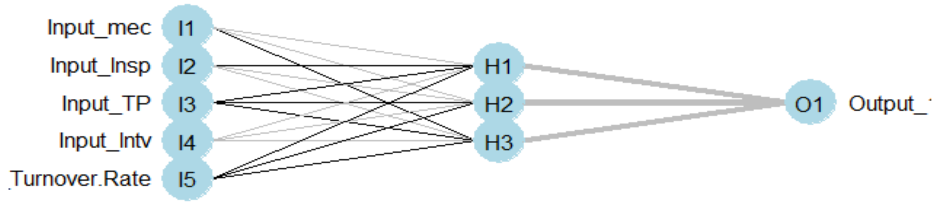

Figure 9: ANN used in the study having one hidden layer and three nodes

The five nodes (I1, I2, I3, I4, and I5) in the input layer took in the five inputs to the model – trading probability (tp), spread (lnsp), Market Efficiency coefficient (mec), trading volume (lntv), Turnover rate, and the output node (O1) provided the computed value of lnIndex. H1, H2, and H3 were the three hidden nodes in the single hidden layer used in the model.

The neural network models were developed in R using the RSNNS library [1], using logistic activation function at the hidden as well as output layers. The performance of the models was evaluated using the Mean Average Percentage Error (MAPE) computed based on the dependent variable computed by the trained model using the data from the respective validation partitions.

As ANN works best with inputs and outputs in the range 0 to 1, we scale the data to that interval while using ANN models. The corresponding output, while using the model for predicting the lnindex was converted back to the original scale for comparison with the observed values and computing the MAPE.

Error measurement statistics play a critical role in tracking forecast accuracy, monitoring for exceptions, and benchmarking your forecasting process. On modeling the liquidity variables using ANN we obtained MAPE of 5.65% for BSE and 5.81% for NSE. MAPE is the relative significance (Percentage) of the error and a value of about 5% using ANN can be considered pretty useful as far as ANN related studies are concerned. Empirical evidences using Normal Regression generally show higher MAPE values.

We had also tried with more complex MLP ANNs, with one to three hidden layers with three to fifteen nodes in each hidden layer, but the best MAPE were obtained for the aforesaid simple network of one hidden layer with three nodes in it. This indicates the presence of a comparatively simpler relationship between the predictor variables and the predicted one.

6.9 Random Forest (RF)

The Random Forest model was built using the Random Forest library of R. The RF was built with 500 trees, and in addition to MAPE for the predicted values of lnindex, relative importance of the different predictor variables was also computed.

As mentioned earlier, while Random Forest is a black box algorithm with a good predictive performance, it does allow certain visibility about the importance of predictor variables used in building the model. One of the important measures that

Random Forest uses to decide on the importance of a given predictor variable is through computation of the increase in MSE of prediction if the given predictor variable has its value permuted, that is, if the values of that predictor variable are replaced with other realistic values. Thus, for a large value of MSEj, the increase in MSE of prediction by permuting the values of predictor j, implies that the predictor j was important in building the model. The MAPE obtained using RF for BSE is 2.17% and for NSE is 2.41%. They are even better than those obtained from ANN models. Both the findings from machine learning tools individually as well as collectively support the existence of a good relationship between the predictor variables (liquidity measures) and the predicted ones (stock market). Figure 10 shows the relative importance of the predictor variables used in the model:

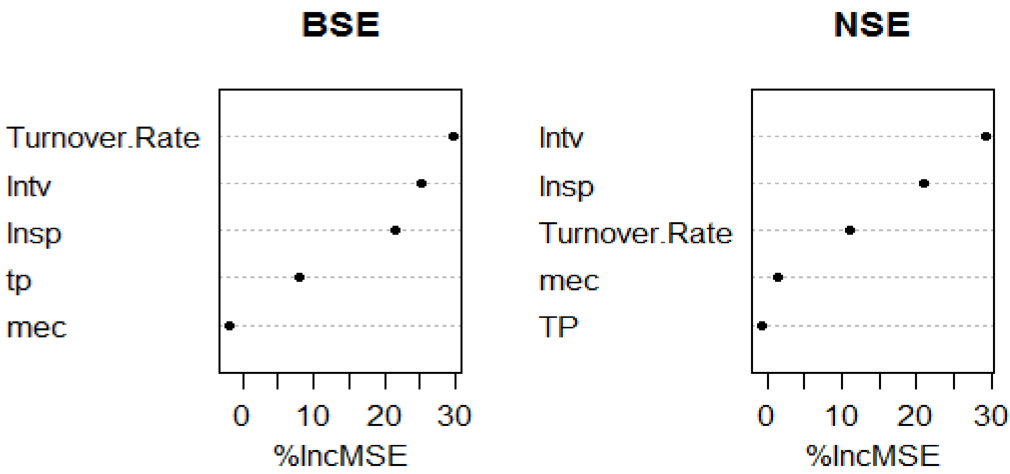

Figure 10: Relative importance of liquidity measures in BSE and NSE

The plot of the predictor variables vis-à-vis the percentage increase in MSE (%IncMSE), the predictor with the highest increase in MSE being the most important player in the model, indicates that the most important predictor variables are Intv (trading volume), Insp (Spread) and Turnover Rate. The findings are consistent for both the exchanges and hence may be generalized for Indian stock market.

6.10 Wilcoxon rank-sum and Kolmogorov-Smirnov tests

The findings are tabulated below in Table 3. It shows the status of null hypothesis of equality (using 5% level of significance) under respective tests for each liquidity measure. Under Wilcoxon rank-sum test, we additionally give an estimate of the probability that a random draw from the first population (i.e., NSE) is larger than a random draw from the second population (i.e., BSE)

We find that liquidity measures in terms of volume, spread, turnover rate and MEC are significantly different in NSE and BSE as per the non-parametric tests used above. The probability estimates that a liquidity parameter of NSE is higher than BSE is more than 50% in all cases except Spread suggesting that there are more chances that BSE might be less liquid than NSE and this finding is consistent with other findings here. Spread is an indicator of tightness and a narrow spread indicates a liquid market. Null hypothesis of equality could not be rejected for trading probability and this may not be surprising as TP is a function of no. of trading (or no of non-trading days) and generally both the exchanges observes holidays on same day in India.

Table 3: Test of Equality between NSE and BSE

	Volume	TR	Spread	MEC	TP
Wilcoxon rank-sum test	Rejected. P{volume(NSE) > volume(BSE)} = 0.725	Rejected. P{TR(NSE) > TR(BSE)} = 0.717	Rejected. P{spread(NSE) > spread(BSE)} = 0.313	Rejected. P{MEC(NSE) > MEC(BSE)} = 0.557	Accepted. P{TP(NSE) > TP(BSE)} = 0.549
Kolmogorov -Smirnov test	Rejected.	Rejected.	Rejected.	Rejected.	Accepted.

7. Conclusion

The paper explores the liquidity position of two broad based stock index from Indian Stock Market in terms of market depth, breadth, and resiliency and attempts to investigate whether the endogenous liquidity measures collectively are capable of explaining changes in those chosen indices. We observe through the time period 2002 to 2015 and under all chosen measures that liquidity was affected during the period of global financial crisis and its recovery period. In fact all measures showed India is still lacking both market depth and breadth when compared to pre-crisis period. The MEC values clearly indicate that resiliency in Indian stock market keeps changing with observed volatility coming down in recent years. The conventional econometric models using time series data show lower levels of accuracy and parameter instability in modeling liquidity and stock market possibly due to non-linearity in the data series. Alternatively, Artificial Neural Networks (ANN) and Random Forest (RF), due to its unique non-parametric, non-assumable, noise-tolerant and adaptive properties, can map any nonlinear function without a priori assumptions and has shown great applicability in time-series analysis and forecasting due to its pattern recognition capability. Using five proxies as a liquidity measures: namely trading probability, spread, Market Efficiency coefficient, trading volume and turnover rate and ANN we obtained a MAPE (mean absolute percentage error) of 5.65% in case of BSE 500 series and a MAPE (mean absolute percentage error) of 5.81% in case of NIFTY 500 series while using RF the errors were lower further. Also RF showed that traded volume, spread and turnover rate (or turnover velocity) are most important liquidity variables for explaining variations in stock market indices. The non-parametric tests indicates that chances are higher that liquidity of the BSE is lower compared to the NSE. This supports the BSE’s latest decision to offer 'Liquidity Enhancement Incentive Programmes Schemes (LEIPS)' to as many as 166 securities exclusively listed on the stock exchange and create a new sub-group named 'XC' group for companies listed exclusively on it. Overall, the study provides more support to liquidity measures as an important factor for explaining variations in stock market especially in the Indian context.

References

Acharya, V.V. and Pederson, L.H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics, 77(2), 375–410.

Amihud, Y. and Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics, 17(2), 223–249.

Amihud, Y., 2002. Illiquidity and stock returns: cross section and time-series effects. Journal of Financial Markets, 5(1), 31-56.

Bali, T. G., Peng, L., Shen, Y and Tang, Y., 2014. Liquidity Shocks and Stock Market Reactions. Review of Financial Studies, 27(5), 1434-1485.

Chordia, T., Roll, R. and Subrahmanyam, A., 2001. Market Liquidity and Trading Activity. The Journal of Finance, 56(2), 501-530.

Constantinides, G. M., 1986. Capital market equilibrium with transaction costs. Journal of Political Economy, 94(4), 842–862.

Datar, V. T., Naik, N. Y. and Radcliffe, R., 1998. Liquidity and stock returns: An alternative test. Journal of Financial Markets, 1(2), 203–219.

Dutta, G., Jha, P., Laha, A. K. and Mohan, N., 2006. Artificial Neural Network Models for Forecasting Stock Price Index in the Bombay Stock Exchange. Journal of Emerging Market Finance, 5(3), 283-295.

Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica, 50(4), 987-1008.

Faff , R., Chang, Y.Y. and Hwang, C.Y., 2010. Liquidity and stock returns in Japan: New evidence. Pacific-Basin Finance Journal, 18(1), 90–115.

Haykin, S. 1999. Neural Networks: A Comprehensive Foundation, NJ: Prentice-Hall.

Heaton, J., & Lucas, D. J., 1996. Evaluating the effects of incomplete markets on risk sharing and asset pricing. Journal of Political Economy, 104(3), 443–487.

Jonathan A. Batten and Xuan Vinh Vo., 2014. Liquidity and Return Relationships in an Emerging Market. Emerging Markets Finance and Trade, M.E. Sharpe, Inc., 50(1), 5-21.

Korajczyk, R. A. and Sadka, R., 2008. Pricing the commonality across alternative measures of liquidity. Journal of Financial Economics, 87(1), 45–72.

Krishnan, R and Mishra, V., 2013. Intraday liquidity patterns in Indian stock market. Journal of Asian Economics, 28, 99-114.

Kumar, G and Misra, A. K., 2015. Closer view at Stock Market Liquidity: A Literature Review. Asian Journal of Finance and Accounting, 7(2), 35-57.

Kumar, G and Misra, A. K., 2015. Analysis of Liquidity on Indian Mid Cap Stocks. International Journal of Economics and Finance, 7(10), 112-125

Marshall, B. R., & Young, M., 2003. Liquidity and stock returns in pure order-driven markets: Evidence from the Australian stock market. International Review of Financial Analysis, 12(2), 173–188.

Martinez, M. A., Nieto, B., Rubio, G., and Tapia, M. 2005. Asset pricing and systematic liquidity risk: An empirical investigation of the Spanish stock market. International Review of Economics and Finance, 14, 81–103.

Moore, L. and Sadka, R. 2006. How fundamental is the pricing of liquidity, Working Paper, Northwestern University.

Narayan, P.K. and Zheng, X., 2011. The relationship between liquidity and returns on the Chinese stock market. Journal of Asian Economics, 22(3), 259-266.

Pastor, L. and Stambaugh, R.F., 2003. Liquidity Risk and Expected Stock Returns. Journal of Political Economy, 111(3), 642-685.

Qin, Q, Wang, Q, Li, J. and Shuzhi, S. G., 2013. Linear and Nonlinear Trading Models with Gradient Boosted Random Forests and Application to Singapore Stock Market. Journal of Intelligent Learning systems and Applications. 5(1), 1-10.

Sala, J.P., 2011. Liquidity Risk Modelling Using Artificial Neural Network. Germany: Lambert Academic Publishing.

Sarr, A. and Lybek, T., 2002. Measuring Liquidity in Financial Markets, IMF Working Paper, No. WP/02/232, 1-63.

Theofilato, K., Likothanassis, S. and Karathanasopoulos, A., 2012. Modeling and Trading the EUR/USD Exchange Rate Using Machine Learning Techniques. ETASR - Engineering, Technology & Applied Science Research, 2(5), 269-272.