Pacific Business Review International
Volume 13 issue 1 July 2020

Determinants of Stock Option Listing: Logistic Regression and Random Forest Approach

Dr. Himanshu Joshi
Associate Professor
FORE School of Management
B-18, Qutub Institutional Area
New Delhi

Dr. Rajneesh Chauhan
Professor
FORE School of Management
B-18, Qutub Institutional Area
New Delhi

Abstract

The study examines the factors affecting stock option listing for Indian market usingbinary logistic regression and machine learning classifier algorithm, random forest. Findings suggest that large size firms paying regular dividends and having high institutional holdings demonstrate higher propensity of stock option listing. On the other hand, firms with high idiosyncratic return variations generally have lower probability of option listing. Results of machine learning algorithm confirm that firm size and idiosyncratic return variations are the two largest influencers of stock option listing, followed by stock volatility, dividend payout, institutional holding, profitability, firm age, leverage, research intensity, employee stock option, and cross listing of firm's stock on multiple exchanges. Overall, besides firm size, any characteristic of the stock which aids in reduction of information asymmetry improves the propensity of stock option listing.

Key Words: Stock Options Listing, Logistic Regression, Machine Learning, Random Forest, Idiosyncratic Return Variation.

Introduction

Financial derivatives were introduced primarily to benefit risk averse investors and managers to hedge their investment and operational risk. Forwards and futures are contracts to buy (or sell) an underlying asset at a predetermined price during the life of the contract or on expiry of the contract. Further, in forwards and futures contract entails that both parties are under an obligation to honor the contract during the life of the contract or on expiry of the contract. By comparison, an option is a contract that gives the parties a right to buy or sell the underlying asset at a predetermined price during the life of the contract or on expiry of contract. Therefore, value of an option depends considerably on the volatility of the underlying asset, in addition to the other factors like expiration time, risk free rate, exercise price, and spot rate. Basically, options on volatile assets, get better pricing in option trading. Theoretically, option payoff can be generated synthetically by constructing a portfolio using the underlying assets and a risk-free security. Then in such theoretically perfect markets, option trading should not impact the risk and return of the underlying asset. However, real markets are incomplete and operates under the frictions of transaction cost and information asymmetry, and in such markets, options listing and trading may influence the underlying assets' risk, return, and trading volume by reducing their information asymmetry. Hence the option market price, volatility, liquidity etc. may get influenced in these imperfect markets.

Option trading began in 1973 on the Chicago Board Options Exchange (CBOE) with contracts on equity indices, exchange traded funds, foreign exchange, interest rates and common shares. Unlike the listing of equity shares, which is sponsored by listing firms, decision to list options are made by the exchanges. World over exchanges set the initial listing requirements that underlying securities must meet in order to list options on them. Largely, shares are selected for option listing by a committee comprising of members of the exchange. As stock exchanges are run with the objective of profit maximization for its members, there is a strong inclination towards listing of options on those shares which are capable of generating highest trading volume. Apart from the profit motive for its members, option listing also get influenced by broader institutional environment in the respective market. For example, in United States, the option exchanges are members of Option Clearing Corporation (OCC) and are subject to federal securities laws and regulatory compliance of the Securities Exchange Commission (SEC). In India, option listing on individual equity shares began in July 2001 on the National Stock Exchange (NSE). Currently, options are traded on 147 individual equity shares confirming the guidelines stipulated by Securities and Exchange Board of India (SEBI). The eligibility requirements stipulated by SEBI for listing an option on the underlying equity shares primarily focus on stock's market capitalization, average daily trading volume, stock's median quarter sigma order size, and average daily deliverables. At the onset, stocks selected by the exchanges for option listing were of large and reputed firms with high trading volume, but later on focus shifted on selecting stocks with higher volatility.

Present study intends to identify the determinants of the option listing on individual stocks in India other than those used by exchange and prescribed by the regulators. Our empirical approach is to select the universe of stocks eligible for option listing, and then use binary logistic regression and machine learning based random-forest method to measure the extent to which the probability of option listing is associated with the factors such as volatility, dividend yield, employee stock option, firm age, firm specific return variation, institutional ownership, firm size, and propensity to engage in R&D activities.

The remainder of the paper is organized as follows. Section 2 discusses the related literature; section 3 develops the relevant hypotheses; The remainder of the paper is organized as follows. Section 2 discusses the related literature; Section 3 develops the relevant hypotheses; Section 4 discusses and delineates the sample, variable measurement and methodology: Section 5 focuses on analysis and findings. Lastly, Section 6 emphasizes on conclusion.

Literature Review

Merton (1973) and Black (1975) postulated that in complete markets, derivative instruments like options and futures are worthless, and their payoffs can be created synthetically using a portfolio of basal assets and a risk free security. Also, it is well established in finance literature that capital markets have information asymmetry and are incomplete. In markets with information asymmetry, Ross (1976) was a pioneer in postulating that options trading can communicate pertinent information by expanding contingencies covered by traded securities. Ross (1976) and Hakansson (1982) showed that options help in making the markets complete. They argued that options provided hedging opportunities for traders. Further, Black (1975) put forward that options can contribute to more informed trading in the underlying assets. Reason being this provided higher leverage to investors who were financially constrained.

Several scholars have studied the effect of option listing on underlying asset. Damodaran and Subrahmanyam (1992), and Mayhew (2001) provide excellent surveys of theoretical and empirical literature on the subject. Most empirical studies reported substantial reduction in stock volatility post option listing, refer Conrad (1989), Detemple and Jorion (1990), and Damodaran and Lim (1991). However, Bollen (1998) reported that on the listing of options on selected stocks, apparent reduction in volatility was recorded even for those stocks having no option listing. This suggest that volatility reduction effect can be spurious. In the context of India, Joshi (2018a) studied the influence of trading in single stock options on volatility of underlying stocks. In this research there was no statistically significant decline in short term volatility or long run volatility. This was found for options of small cap and mid cap firms.

Another tranche of literature acknowledges that option trading makes underlying stock market more efficient by inducing informed trading. Figlewski and Webb (1993) and Johnson and So (1992) found that option trading supports trading that is more informed. This is a consequence of relaxing the short-sale constraint on the underlying asset. Also, Cao (1999) reported that listing and trading of options, propels traders with comparatively less market information, to gather private information and knowledge regarding the underlying asset. Chakravarty, Gulen and Mayhew (2004); and Pan and Poteshman (2006) have argued that such private knowledge and information is very useful for investments that have a long term outlook. Such information contributes to making the stock market relatively more efficient.

Options trading plays a very important role by decreasing information asymmetry in the market and thereby completes the market on account of reasons like higher leverage opportunity provided to traders who are financially constrained but informed, by lifting short sale constraints on stocks and pushing traders to search for more private knowledge and information about the stocks in question. Further as postulated by (Ross 1976) and Hakansson (1982), avenues of hedging opportunities open up and this in turn also leads to more trading demand of the underlying stocks. Hedging transactions in incomplete markets replete with information asymmetry reduces the chances of uninformed market transactions. Black (1975) postulated that since options provides opportunities of leverage to informed investors hence informed trading increases in the market. Easley, O'Hara and Srinivas (1998) put forward that between informed investors and uninformed informed, informed investors find options more attractive as they find availability of complex and multiple contracts less daunting.

Further, researchers have established a link between trading in options and the underlying asset volatility, price etc. In some researches it has been established that trading in options gives the traders information about price volatility of the stock price (Ni, Pan, and Poteshman, 2008). In other related researches by Chakravarty, Gulen and Mayhew (2004) and Pan and Poteshman (2004) it has also been opined that volumes of options traded indicate the likely direction of the price of underlying stock

If stock markets are more efficient, then traders with less information make a conscious effort to know about the fundamentals of the firm. This action in turn reduces problems of information asymmetry. This is especially helpful when traders are evaluating firm's long term investment, like Capex investments, R & D investments etc. In a study by Blanco and Wehrheim (2017) it has been established that less information asymmetry as a consequence of options leads to innovativeness effort of firm.They argue that for firms that are listed on options markets, greater trading activity is associated with an increased propensity to innovate. Similar study has been conducted by Joshi (2018 b) in the context of India. Author examined the effect of option trading on firm-level innovation for publicly listed Indian firms. He found that the firm profitability, past financial leverage, dividend payout ratios over the years, and the age of firm, and affect the innovativeness of the firm.

Literature on empirical determinants of stock option listing are scant. Cowan, Carter, Dark, and Singh (1992) studied the empirical determinants of equity stock listing on the New York Stock Exchange (NYSE). Closest prior work to the present study has been conducted by Mayhew and Mihov (2004). They studied listing choices made by the option exchanges in the US, and found that exchanges tend to list options on stocks with high trading volume, volatility, and market capitalization. No similar study has been conducted in context of the emerging market.This study is an attempt to establish determinants of stock option listing in an emerging market. Similar studies have been done in developed market but not in emerging markets. Given the inherent differences between emerging and developed markets in terms of information asymmetry this study is different and novel. Moreover, we have used classifier machine learning algorithm namely, random forest to confirm the determinants of option listing predicted by binary logistic regression. We have advanced the work of Mayhew and Mihov (2004) by including additional explanatory factors for stock option listing. In addition to the logit framework used by Mayhew and Mihov (2004), we have applied machine learning based algorithm namely, Random Forest. Our work focuses on the stock option listing in the Indian market. Random Forest is based on the notion of bootstrap aggregation, which is method for resampling with replacement in order to reduce variance.

Leo Breiman came up with the concept of Random Forests, a concept that improved accuracy of Decision Trees and builds on bagging of decision tree. The concept of Random Forests was influenced by an earlier work done by Amit and Geman (1997), where random selection of geometric features was done for best split at each node (Breiman, 2001). Likewise, Random Forests builds on bagging. Bagging predictors is a method of generating multiple versions of a predictor and using these to get an aggregated predictor. Leo Breiman's seminal paper named Random Forests in 2001, encapsulates and articulates the concept of Random Forest very comprehensively. Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges to a limit as the number of trees in the forest becomes large. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them (Breiman, 2001). Random Forests are an improvement over bagged trees because random forest because de-correlates the trees.

When the decision trees are built a random number of predictors are used to as candidates for splitting. This way the trees formed are not so correlated. Random Forests come across as an effective ensemble machine learning method. Random Forests can be used for regression as well as classification. When used for classification, a random forest obtains a class vote from each tree, and then classifies using majority vote. Applications of Random Forests are numerous and only a few can be mentioned here. In the field of finance, they have been used to forecast high growth companies (Weinblat, 2018), corporate governance risk (Creamer and Freund, 2004), financial fraud detection (Liu et. al., 2015), trading strategies for futures (Cheng and Chiang, 2019) and many more. Thus, we see that applications of random forest have met success across varied fields and therefore, Random Forest has been used as a classifier in this research. Random Forests algorithm has been implemented using random Forest package in R software.

Hypotheses Development

Traditionally exchanges have used trading volume, volatility, and firm size as the primary criterion for listing of options on the stocks. Exchanges chose stocks with high trading volume because high trading activity in the underlying asset will induce higher trading in the option contracts as well, which in turn will be profitable for the exchanges. Since, pricing of an option is substantially influenced by the underlying asset's volatility, hence a stock that exhibits volatility in its price will have more chances that the stock's option contracts gets listed. Generally, exchanges list options on stocks of large and well known firms, which is again related to the trading volume. Large and reputed firms are part of various national and international indices, and both active and passive fund managers hold these stocks in their portfolios, which generate higher trading volume for these stocks.

In the present study, we propose certain additional factors that can influence probability of option listing on the stocks. Institutional investors are hypothetically more informed than the retail investors. They collect and process public information about the firm in sophisticated manner than the ordinary retail investors. Also, due to consolidated shareholding, institutional investors may also influence strategic decision making of the firm. Thus an institutional investor, who is privy to strategic decisions of the firm is better placed to gain from trading of stock options of that firm. Therefore, we propose that higher institutional holding can encourage exchanges to list options on such stocks.

Dividend payout and cross listing dummy are surrogates for higher information symmetry and also more open and transparent disclosures. Generally, dividend payout is a matter of firm's financial policies such as reinvestment opportunities, cash holdings, and clientele shareholders. However, in emerging markets firms use dividend payments as signal of upright disclosure of earnings. A firm that reports good earnings and simultaneously announced attractive dividend payments, confirms the quality of its disclosure by ensuring that it has sufficient cash payout dividends. Similarly, cross listing of stocks on exchanges of developed markets where corporate governance and disclosure norms are more stringent, ensures lower information asymmetry. Therefore, we hypothesize that firms that pay higher dividends and have cross listed their stocks on international exchanges, their stocks have better prospects of stock option listing. A dummy variable has been included to represent whether the firm has Employee Stock Options (ESOP)s given to employees.

The ratio of firm's R&D expenditure to firm's total assets is a manifestation of R & D intensity. A firm having high intensity of research is likely to possess plenty of project specific technical information that is difficult to interpret for the outsiders. Generally, specialist investors who hold expertise in analyzing such projects, trade in such stocks. Blanco and Wehrheim (2017) argue that stock option trading on such R&D intense stocks induces informed trading. To compete with informed traders, uninformed traders gather more information about research activities of the firms, which result into reduced information asymmetry. They argue that for firms that are listed on options markets, greater option trading activity is associated with an increased propensity to innovate. Therefore, we hypothesize that stocks of the R&D intensive firms have better prospects of option listing.

The variable Ψ measures firm-specific idiosyncratic stock return variation relative to market-wide variation, or lack of synchronicity with the market. French and Roll (1986) and Roll (1988) postulated that idiosyncratic variations in firm specific return, indicates information asymmetry and private information. All things remaining the same, more is the variation in firm specific returns, more is information asymmetry and private information. So, our hypothesis in case of Ψ is that stocks having higher value of Ψ have lesser odds of stock option listing. Another explanatory variable for stock option listing considered in the study is firm age. Apparently, relationship of firm age with probability of stock option listing is paradoxical. Generally, exchanges select large, reputed firms with high trading volume for option listing, which make well established firms a good fit for stock option listing. On the other hand, new technology firms which are intensely engaged in research and development endeavors and possess loads of specialized and difficult to interpret information are likely to benefit from the stock option listing.

Data and Methodology

Options are listed on 147 stocks on national stock exchange of India (NSE). NSE follows guidelines of the Securities and Exchange Board of India (SEBI) for listing option contracts on the stocks. The stock on which option is to be listed, must be chosen from among the largest 500 stocks in terms of average daily market capitalization, and average daily traded value.

Firm specific cross sectional data for the financial year 2018-19 was collated for the largest five hundred listed Indian firms (in terms of market capitalization) from Thomson Reuters Eikon database. The variables under consideration were many and after excluding the companies for which the data was not available for all the companies, 208 listed companies were left. Next, in this set of 208 companies, it was only 89 firms for which option trading was active. To elicit the determinants of option listing on stocks, binary logistic regression analysis has been used. Dependent variable is a dummy variable signifying option listing. Independent variables used in the study are volatility of stock returns, dividend payout ratio, dummy for cross listing, firm age in years, measure of firm specific return variation, institutional holding, firm size, and R&D intensity are specified as independent variables.

Option Listing Dummy = $\beta0$ + $\beta1$ (Volatility) + $\beta2$(Dividend Payout) + $\beta3$ (Cross Listing Dummy) + $\beta4$(Firm Age) + $\beta5$ (Firm Specific Return Variation) + $\beta6$(Institutional Holding) + $\beta7$ (Firm Size) + $\beta8$ (R&D Intensity) + εi. [1]

Idiosyncratic return variation is measured running regression under capital asset pricing model. Since, R2 of the regression estimates the return on the particular stock explained by the market return, firm specific return variation is estimated by 1- R2. Given the bounded nature of R2, a logistic transformation has been computed as follows:

$\Psi = \text{Ln}\,[(1\text{-}R2)/R2]$[2]

The variable Ψ measures firm-specific idiosyncratic variation relative to market-wide variation, or lack of synchronicity with the market.

In addition to the binary logistic regression method, a machine learning algorithm namely Random Forest has been used in the study. Random Forests are an effective tool of predicting outcomes. They give results which are competitive with other methods and their prediction accuracy is better, they reduce bias, are robust to noise. Random inputs and random features produce good results in regression and classification—and especially more so in classification.

Random Forests have an advantage that there are a very few tuning parameters. There are only two main tuning parameters namely the “number of trees” and the “number of predictors” used while making decision trees. Number of Trees (ntree): A Random Forest is a forest of trees where number of trees are built. When used for classification, a random forest obtains a class vote from each tree, and then classifies using majority vote. By default, the 'number of trees' are 500. Prudence demands that when making Random Forests different 'number of trees' should be tried out and the corresponding error rate should be seen. In this research, the “number of trees” that lead to minimum error rate was 2000. Hence the value of the tuning parameter “number of trees” is 2000.

Number of Predictors (mtry): When making a tree, rather than using all the predictors, even the predictors can be selected randomly. The 'number of predictors' used as candidates for making decision trees is a tuning parameter. For classification problems, the default value of “number of predictors” is $\sqrt{}$ p and the range in which “number of predictors” will vary is between a minimum of 1 and maximum of p, where p is the maximum number of predictors. For different values of “number of predictors” the corresponding error rate is tabulated. In this research, the “number of predictors” that lead to minimum error rate was 3. Hence the value of this tuning parameter “number of predictors” is 3.

Findings and Analysis

The summary statistics of independent variables used in the study are provided in the below mentioned table i.e. Table 1. This Table has been segregated into two parts, Part (A) and Part (B) to provide the summary statistics separately for the companies with option trading and companies with no option trading.

Table 1. Summary Statistics for Independent Variables (without option trading and with active option trading).

A. Without Active Option Trading (119 Firms)	Mean	Median	Std. Dev	Min	Max
Firm age	25.728	26.449	10.290	1.940	36.934
Institutional Ownership	0.430	0.425	0.157	0.100	1.000
Dividend Payout	17.568	10.050	25.693	0.000	159.150
Research Intensity	0.010	0.004	0.018	0.000	0.148
Financial Leverage	0.393	0.140	0.693	0.000	6.140
Cross Listing	0.084	0.000	0.279	0.000	1.000
Firm Specific Return Variation	0.303	0.128	0.557	0.001	4.807
Volatility	2.103	2.030	0.515	0.950	3.940
Firm Size	8.629	8.537	0.800	6.861	11.073
B. With Active Option Trading (89 Firms)	**Mean**	**Median**	**Std. Dev**	**Min**	**Max**
Firm age	28.150	25.222	14.235	1.841	80.164
Institutional Ownership	0.500	0.479	0.167	0.250	1.000
Dividend Payout	35.081	28.440	32.416	0.000	178.320
Research Intensity	0.013	0.004	0.021	0.000	0.128
Financial Leverage	0.374	0.110	0.640	0.000	3.930
Cross Listing	0.236	0.000	0.427	0.000	1.000
Firm Specific Return Variation	0.032	0.014	0.062	0.000	0.473
Volatility	1.796	1.740	0.413	1.110	3.760
Firm Size	10.323	10.462	1.237	7.505	13.080

Source: Table constructed by the author using firm level data from Thomson Reuters' Eikon Database.

In comparison to the non-option trading firms, option trading firms are larger in size, older in terms of firm age, more liquid, more engaged in research and development activities, greater institutional ownership, and have higher propensity of cross listing and dividend payments. On the other hand, non-option trading firms aremarginally more leveraged and volatile.Apparently, it can be inferred from the table 1 that higher intensity of research and development activities leads to higher probability of stock option listing.In conjunction with cross listing, stock option trading also moderates information asymmetry and in turn induces informed trading. These firms are more likely to offer employee stock options which help restrain the agency cost associated withowner-management conflict.Reduced information asymmetry and restrained agency cost manifest in substantially lower idiosyncratic return variation for option-trading firms.

The results of binary-logit regression analysis are presented in the below mentioned table, i.e. Table 2. In this analysis, dependent variable was option trading dummy while the independent variables are volatility of stock returns, dividend payout ratio, the age of firm, institutional holding, firm size, firm specific return variation, ratio of R&D expenditure to total assets, and cross listing dummy.

Table 2. Binary-Logistic Regression Analysis Results on Option Trading Dummy.

	Binary Logit (Quadratic hill climbing) ***Option Trading Dummy***
McFadden R-squared	*0.531*
LR-Statistic *Prob (Wald F -Statistics)*	*149.979****
Return Volatility	*-0.192* *-0.371*
Dividend Payout	*0.0171* *1.981***
Firm Age	*-0.027* *-1.369*
Institutional Ownership	*5.711* *3.306****
Firm Size	*1.499* *5.013****
Firm Specific Stock Return Variation (Ψ)	*-7900722* *-2.399****
Research Intensity	*12.689* *1.0301*
Cross Listing Dummy	*1.051* *1.809****

(*p<0.10, **p<0.05, ***p<0.01.)

In binary-logistic regression, coefficients of explanatory variables validateprospects of stock option listing. Coefficients of dividend payout, institutional ownership, firm size, idiosyncratic return variation, and cross listed dummy have statistically significant coefficients. Out of these statistically significant variables, dividend payout, institutional ownership, firm size, and cross listed dummy have positive coefficients, whereas, idiosyncratic return variation (Ψ) has negative coefficient. Positive dividendpayout coefficientconfirms that dividend paying mature firms which send out signals to the investors that quality of disclosure on earnings is upheld, have better prospects for stock option listing. Similarly, firms having higher institutional holding tend to enjoy more informed trading, and prospects stock option listing seems bright for these firms. Positive coefficient of firm size also endorses the empirical evidence from earlier studies in developed markets that exchanges prefer to list stock options on larger firms. Positive coefficient of cross listing indicates towards better compliance and lower information asymmetry, which improve the prospects of stock option listing. Idiosyncratic return variation has a negative coefficient, showingthat firms whose returns are not aligned with the market returns are unlikely to attract stock option listing.Remarkably, coefficient of return volatility has a negative value but it is not statistically significant. This is contrary to the earlier evidence provided by Mayhew and Mihov (2004) in the context of developed market that over the years, volatility has become an important determinant of stock option listing. When we try to infer from the combined results of firm specific return variation and return volatility, former being statistically significant, and later statistically insignificant, it is evident that for stock option listing synchronization of stock return with market return is important, but not the indigenous volatility. Positivecoefficient of research intensitypoint towards the higher probability of option listing for these firms. However,as thecoefficient of R&D intensity is not statistically significant, result is merely indicative.Similarly, coefficient of firm age is negative but statistically insignificant, indicating relative advantage of newer firms over old firm in terms of stock option listing.

French and Roll (1986) and Roll (1988) postulated that inexplicable variations in firm specific return, indicates information asymmetry and private information. All things remaining the same, more is the variation in firm specific returns, more is information asymmetry and private information.

Stocks of such firms have low chances of getting option listed. Having said that this also is a fact that stocks of such firms when option listed can lead to aggregation of information and diffusion of information on account of trading in options. This in turn leads to less of information asymmetry and efficient stock process. The results ratify the assumption that traders will gain from information generated by trading of options about the likely direction of underlying stock prices.

To improve the prediction accuracy of binary logistic regression model for stock option listing, we have used a machine learning based algorithm namely, random forest. Since random forest algorithm for classifier problem works very well with large number of explanatory variables, we have added five new explanatory variables in addition to the list used for binary logistic regression. These additional explanatory variables are return on assets, annual R&D allocation, a dummy for employee stock option (ESOP), EPS growth rate, and financial leverage. Return on assets measures firm's overall profitability; dummy for employee stock option captures whether firm has offered stock options to its key employees or not. ESOP dummy has been used as a proxy for reduced agency cost. Firm leverage has been calculated as total debt divided by total assets. Annual R&D allocation has been calculated as annual R&D expense divided by total revenue for the year. Table 3 presents the mean decrease in Gini for explanatory variables. More is the decrease in Gini coefficient, more important is the variable.

Table 3. Explanatory Vari ables and Mean Decrease in Gini Coefficient.

Explanatory Variables	Mean Decrease in Gini
Firm Size	28.240
Firm Specific Return Variation (Ψ)	22.500
Volatility	8.624
Dividend Payout	8.369
Institutional Holding	6.887
Return on Assets	6.574
Firm Age	5.530
Leverage	4.124
Annual R&D Allocation	3.919
Overall R&D Intensity	3.867
ESOP Dummy	1.350
Cross Listing Dummy	1.221
EPS Growth Rate	0

Random forest algorithm results also confirm that the most important measures for stock option listing are"firm size", and "firm specific return variation". While exchanges worldwide consistently use the firm size as an important criterion for stock option listing, firm specific return variation is the contribution of our study. As per the results of binary logistic regression, coefficient of idiosyncratic return variation was negative, indicative offirms having more market-dependent return variation, have higher propensity of stock option listing. High firm specific return variation can be construed as a sign of stock illiquidity. Therefore, in simple terms we can endorse that more liquid firms have higher predisposition for stock option listing. This result is in contrast with earlier studies conducted in developed markets which reported stock volatility as second most important criteria for stock option listing after firm size. This is because in emerging markets where liquidity is relatively low, exchanges prefer to list options on the stocks which are well aligned with the market returns. Third most important factor for stock option is listing is stock volatility, followed by dividend payout, and institutional holding. Firm leverage is another factor that emerges out to be an influencer of stock option listing. Interestingly, out of the two parameters taken to denote firms' research and development initiatives namely, "annual R&D allocation" and "overall R&D intensity", the former has slightly better influence on the stock option listing. Dummy variables of ESOP and cross listing have very small contraction in Gini coefficient, demonstrating very little influence of these variables on propensity of stock option listing. Finally, there was zero decrease in the Gini coefficient for EPS growth rate, indicating no influence of it on the stock option listing. EPS growth rate which is a key determinant of stock price, was added to validate any influence of such stock based influencers on stock option listing. Figure 1 shows the mean decrease in Gini coefficient for all the explanatory variables of stock option listing.

Figure 1 : Mean Decrease in Gini

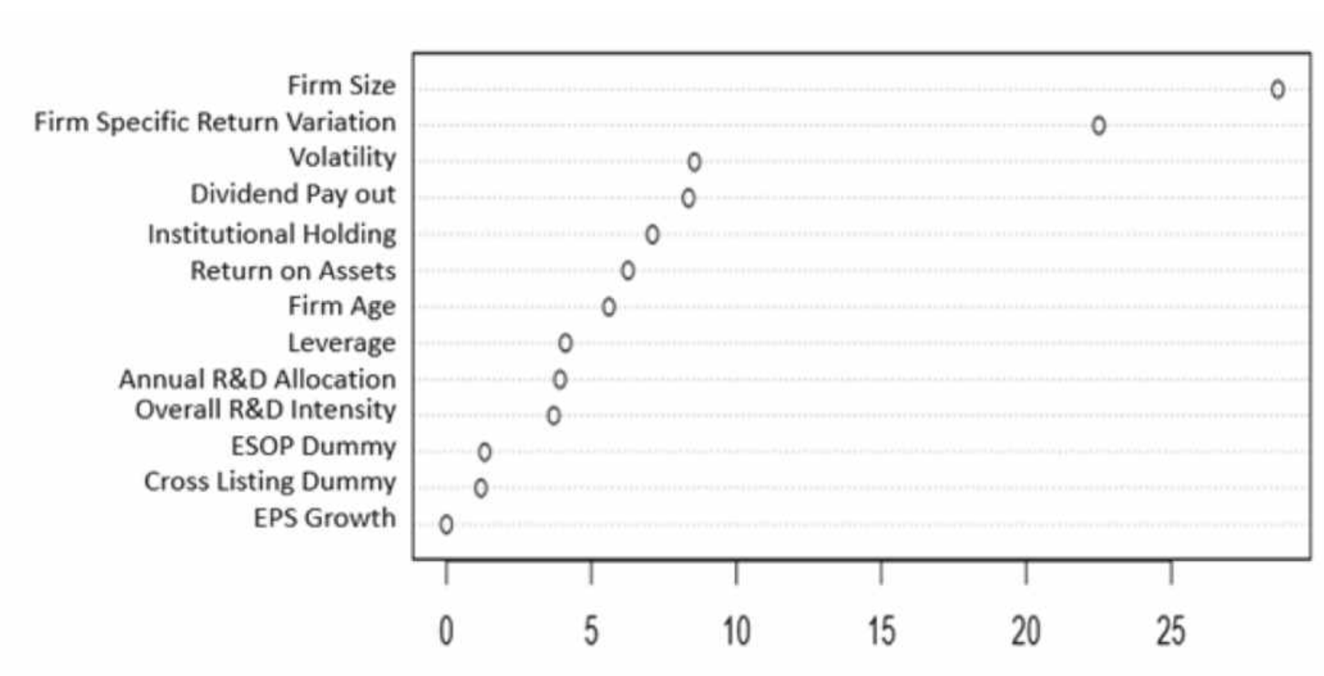

Table 4 presents the result of random forest algorithm in contingency matrix. Based on the contingency matrix generated by random forest algorithm, prediction accuracy of the model for stock option listing has been calculated in table 5.

Table 4. Contingency Matrix for Stock Option Listing using Random Forest Algorithm.

		Predicted		
		Listing	No Listing	Class Error Rates
Actual	Listing	72	17	19.10 %
	No Listing	10	109	8.4 %

Table 5. Prediction Accuracy for Stock Option Listing Model Generated by Random Forest Algorithm.

Prediction Accuracy Parameter	Calculation Formula	Prediction Accuracy Value
Sensitivity or True Positive Rate	$\frac{True\ Positive}{True\ Positive + False\ Negative}$	0.8089
Specificity or True Negative Rate	$\frac{True\ Negative}{True\ Negative + False\ Positive}$	0.9159
Total Accuracy	$\frac{True\ Positive + True\ Negative}{Total}$	0.8701
Model Precision	$\frac{True\ Positive}{True\ Positive + False\ Positive}$	0.8780

The overall accuracy of the model was 87.109 %. The Specificity or the True Negative Rate was 0.915 and Sensitivity or the True Positive Rate was 0.8089. Going by the Class Error Rates the prediction error rate for No-Listing is 8.4 % and for Listing it was 19.10 %. Looking at these numbers and the associated accuracy of 87 %, it can be concluded that the classifier model that has so evolved is a good and robust model.

Conclusion and Scope of Further Research

Under binary logistic regression framework, firm size, institutional ownership,idiosyncratic return variation, dividend payout, and cross listing status of the stock influence the probability of stock option listing in Indian market. These factors are arranged in descending order based on the size of their coefficient in the regression estimates. Larger firms which pay regular dividends with higher institutional ownership are the probable candidates for stock option listing. This is because, exchange selects select stocks for option listing on basis of firm size. Regular dividend payments and cross listing of stocks in multiple exchanges are likely to reduce information asymmetry associated with stocks, and option listing on such stocks motivates informed trading. On the other hand, firms with high idiosyncratic return variations are generally have lower probability of stock option listing.This is because, stocks whose returns are not aligned with market returns are likely to suffer from illiquidity, and exchanges would have no incentive to offer any derivative contract in general and options in particular on such stocks. Results of machine learning algorithm- random forest confirm that firm size and firm specific return variations are the two largest influencers of stock option listing, followed by stock volatility, dividend payout, institutional holding, profitability, firm age, leverage, research intensity, employee stock option, and cross listing of firm's stock on international exchanges. Overall, besides firm size, which is regular selection criteria used the exchanges, any characteristic of the stock which aids in reduction of information asymmetry improves the propensity of stock option listing. The present study has used classifier model for determining probability of stock option listing, further research can be taken up considering option trading volume data to augment the empirical evidence of the present research. There is scope of further research in allied fields around option trading in emerging markets. There is scope for going beyond India into other emerging markets and also studying the industry specific vagaries if any.

Acknowledgement

The infrastructural support provided by FORE School of Management, New Delhi in completing this paper is gratefully acknowledged.

References

Amit, Y., and Geman, D. (1997). Shape quantization and recognition with randomized trees. Neural Computations, 9 (7), 1545-1588.

Black, F. (1975). Fact and fantasy in the use of options, Financial Analysts Journal, 31, 36-72.

Blanco, I., and Wehrheim, D. (2017). The bright side of financial derivatives: options trading and firm innovation. Journal of Financial Economics, 125 (1), 99-119.

Bollen, N. (1998). A note on the impact of options on stock return volatility, Journal of Banking and Finance, 22, 1181-1191.

Breiman, L. (2001). Random Forests, Machine Learnings, 45, 5-32.

Cao, H. (1999). The effect of derivative assets on information acquisition and price behavior in a rational expectations equilibrium. Review of Financial Studies, 12 (1). 131-163.

Chang, J., and Chiang, M. (2019). A random walk down the random forest: Ensemble learning assisted trading strategies TAIEX futures. Academia Economic Papers, Taipei, 47 (3), 395-448.

Chakravarty, S., Gulen, H.., and Mayhew, S. (2004). Informed trading in stock and option market. The Journal of Finance, 59 (3), 1235-1257.

Chen, Q., Goldstein, I., and Jiang, W. (2007). Price informativeness and investment sensitivity to stock price. Review of Financial Studies, 20 (3), 619-650.

Conrad, J. (1989). The price effect of option introduction. Journal of Finance, 44,487-498.

Cowan, R., Carter, R., Dark, F., and Singh, A. (1992). Explaining the NYSE listing choices of Nasdaq firms. Financial Management, 21, 73-86.

Creamer, G., and Freund, Y. (2004). Predicting performance and quantifying corporate governance risk for Latin American ADRs and banks. Financial Engineering and Applications. MIT Press, Cambridge.

Damodaran, A., and Lim, J. (1991). The effects of option listing on the underlying stocks' return processes, Journal of Banking and Finance, 15, 647-664.

Damodaran, A., and Subrahmanyam, M. (1992).The effects of derivative securities on the markets for the underlying assets in the United States: A survey. Financial Markets and Instruments, 1, 1-22.

Detemple, J., and Jorion, P. (1990). Option listing and stock returns, Journal of Banking and Finance, 14, 781-801.

Easley, D., O'Hara, M., and Srinivas, P. (1998). Option volume and stock prices: Evidence on where informed traders trade. 53 (2), 431-465.

Figlewski, S., and Webb, G. (1993). Options, short sales, and market completeness. TheJournal of Finance, 48 (2), 761-778.

Ferreira, D., Ferreira, M., & Raposo, C. (2011). Board structure and price informativeness. Journal of Financial Economics. 99 (3). 523-545.

French, K., and Roll, R. (1988). Stock returns variances: The arrival of information and the reaction of the traders. Journal of Financial Economics. 17, 5-26.

Hakansson, H. (1982). Changes in financial market:

Welfare and price effects and the basic theorems of value conversion. The Journal of Finance. 37, 977-1004.

Johnson, L., and So, E. (2012). The option to stock volume ratio and future returns. Journal of Financial Economics. 106, 262-286.

Joshi, H. (2018 a). Does introduction of stock options impact stock volatilityﬁ Empirical evidence from underlying stocks in Indian Market. Theoretical Economics Letters, 8 (10), 1803- 1815.

Joshi, H. (2018 b). Option trading, information asymmetry and firm innovativeness: Evidence from stock option trading firm in India. Theoretical Economics Letters, 8 (11), 2169-2181.

Liu, C., Chan, Y., Kazmi, S., and Hao, F. (2015). Financial fraud detection model based on random forest. International Journal of Economics and Finance, 7 (7),178-188.

Mayhew, S. (2001). The impact of derivatives on cash markets: What we have learnedﬁ Working paper, University of Georgia.

Mayhew, S., and Mihov, V. (2004). How do exchanges select stocks for option listingﬁ The Journal of Finance, LIX (1), 447-471.

Merton, R. (1973). An intertemporal capital asset pricing model. Econometrica. 41(5), 867-887.

Ni, S., Pan, J., and Poteshman, A. (2008). Volatility information trading in the option market. The Journal of Finance. 63 (3). 1059-1091.

Pan, J., and Poteshman, A. (2004). The information of the option volume for future stock prices. NBER Working Paper No. 10925. http://www.nber.org/papers/w10925

Roll, R. (1988). R2. The Journal of Finance. 43, 541-566.

Ross, S. (1976). Options and efficiency. The Quarterly Journal of Economics, 90 (1), 75-89.

Weinblat, J. (2018). Forecasting European high-growth firms – A random forest approach. Journal of Industry, Competition and Trade, 18, 253-294.