Developing Models to Predict Constituent Changes (exclusion/ inclusion) of Securities in NIFTY 50 Index for Maximizing Investment Returns

Satyam Khatri*, Shubhra Johri**

Abstract

A Common Observation regarding stock price behavior pre and post inclusion and exclusion of a company into indices is that share prices tend to increase prior to inclusion and decrease prior to exclusion. Although Eligibility criteria for inclusion and exclusion include Impact Cost (Liquidity), Market Capitalization, Domicile, Eligible types of Securities and Free Float but still if we can devise models that predict inclusion or exclusion using certain key performance indicator that are widely available, then, the stock market can surely provide a lot of scope for making money for the equity investors. This research work aims at developing two such models .Model –1 predicts with extreme accuracy whether or not any company from the existing 50 companies in Nifty Index will be excluded in the near future and Model–2 predicts if a company will be included as Nifty constituent using discriminant analysis. The KPIs used for the purpose of study include Rate of growth of Profit after tax (ROG-PAT), Rate of growth of Net Sales (ROG-SALES), Rate of growth of Market Capitalization (ROG-MCAP) and Interest Coverage Ratio (ICR). The model proposes a range of outcome values which can be used to predict inclusion or exclusion from NIFTY. A long (buy) position if taken in the securities that are predicted by the model to be included in the index, will help in generating higher returns as its financial performance will improve before the inclusion leading to higher stock price. Alternately, a short (sell) position taken in the securities that are predicted by the model to be excluded from the index , will also be a good bet as its financial performance will degrade before the exclusion, leading to lower stock price.

Key words : *Inclusion, Exclusion Nifty, KPIs, long-term and short-term*

Introduction

S&P CNX Nifty, a well-diversified stock index accounting for 22 sectors of the economy includes 50 of the approximately 1430 companies listed on the NSE and captures approximately 67% of its market capitalization and is a true reflection of the Indian stock market. The S&P CNX Nifty owned and managed by India Index Services and Products Ltd (IISL) is a joint venture between the NSE and CRISIL. IISL has a licensing and marketing agreement with Standard and Poor's (S&P).

Significance of the study:

Inclusion and exclusion of securities has always been one of the favorite topics for researchers and .

*Mr. Satyam Khatri,
Asst. Professor, MBA department
G.L.Bajaj, Greater Noida

**Ms. Shubhra Johri,
Asst. Professor, MBA department
ABES Engineering College, Ghaziabad

Although a lot of work has already been done on the post inclusion and exclusion performance of stocks but when it comes to prediction whether a company will be included or excluded from the index, still remains unexplored and offers a lot of scope for further research and analysis. If we go by books, academicians and general practices, some specific criteria have been stated for selection of Nifty 50 Constituent Stocks viz. Impact Cost (Liquidity), Market Capitalization, Domicile, Eligible types of Securities and Free Float. The deciding factors are that the security should have been traded at an average impact cost of 0.50% or less during the last six months, must have at least twice the market capitalization of the current smallest index constituent, must be domiciled in India and traded on the NSE, should be an eligible security, should have at least 10% of its stock available to investors (float).

But predicting inclusion or exclusion on the basis of these factors is a real tough task for a normal investor. Needless to say that had there been an easier method to be used by a normal investor for such predictions, it could prove to be of great value for them. It is a general observation that financial performance of a company is seen improving before its inclusion in index and is seen degraded before its exclusion. Such specific observable patterns if tamed constructively can offer a lot of scope for making money. A long (buy) position if taken in the securities that are going to be included in the index, shall generate high returns as its financial performance will improve before the inclusion and stock price will increase as well. Alternately a short (sell) position, if taken in the securities, are going to be excluded from the index, will even generate good returns as the financial performance of such securities shall degrade before the exclusion leading to lower stock prices.

But the trick of the trade is indeed an accurate prediction of inclusion and exclusion on part of the investor. This paper is an attempt to develop models which can be used to predict inclusion or exclusion of companies using widely available key performance indicators such as Profit after Tax (PAT), Sales, Market Capitalization and Interest coverage ratio (ICR).

Motivation of Study

Every equity investor aims at maximizing returns and reducing the risk. They are constantly in search of ways for safe investment. If we talk about risk – return profiles, then derivatives have almost highest risk return profile, equities are moderate in this profile and mutual funds are last in the queue. Mostly our focus is on investments in equities. A real investor is a person who can exploit the present opportunities available in the market to maximize the returns and minimize the risk. An investor who stays update with the market regularly often observes some trends and patterns in the market and uses them for investment decision making. This research model is based on one such observation of the pattern that will help to maximize returns and minimize risks.

Literature Review

Over years, a substantial number of research has been done on stock market inclusion and exclusion. Amongst them, a number of researchers have made some really sincere efforts. Shleifer (1986) studied inclusions to the S&P 500 over the period from 1966 to 1983 and concluded that there are significant positive abnormal returns as soon as announcements are made regarding inclusion. Jain (1987) found the mean excess return on the announcement day amounting to 3.07% using the market model. To further this research, Harris and Gurel (1986) concluded that buying the newly included stocks the next day following the announcement and selling them at a higher price afterwards could have helped make great returns for period covering 1973 - 1983. Findings of Beneish and Whaley (1996) and Lynch and Mendenhall (1997) lay emphasis on the fact that the total price effects following addition to the S&P500 index since 1989 reach 7.2%. A number of researchers were of the opinion that trading volume and trading size increases in case of inclusion and also there is a sizeable decrease in the quoted bid/ask spread of the newly included stocks. Many of the studies even acknowledge changes in the trading volume post inclusion. As per the information gathered by Howard and Chan (2002), a number of studies have supported the fact that there are significant price effects associated with changes in the composition of market indices, particularly the S&P 500. Over the period 1976 to 1988 when Standard and Poor's announced and implemented changes in the index sample simultaneously, additions were

associated with an average abnormal return of approximately 3% on the first trading day after the change. The majority of studies found that the price changes were sustained over subsequent trading days. Since October 1989, Standard and Poor's had generally announced index changes a week in advance, the price response was larger. The results of these US studies appear to be robust to variations in the methodology used. In addition, a number of researchers have recognized abnormal trading activity after the announcement of index changes. Masse et al (2000) found that stock prices respond positively to inclusion, outperforming market by 4.29%. They even stressed the fact that stock prices start adjusting before any announcement of inclusion is made. In Europe, price adjustments associated with addition to an index have been examined in the UK. Brealey (2000) reported that stocks added to both the FTSE-All Share and the FTSE 100 indices experienced, on average, a positive abnormal return over an 11-day period that included both the announcement and the effective day. The first study of stock inclusions in indices carried out in emerging markets was that of Hacibedel and Bommel in 2006 in which returns of emerging market stocks were included in the MSCI Emerging Markets. A positive permanent price change amounting to 2-3% upon index inclusion was observed.

Research Methodology : Data Sources

Data on previous inclusion and exclusion of securities from Nifty Index have been taken from National Stock Exchange, India website www.nseindia.com. Data on Key performance Indicators (KPIs) of the companies such as Profit after Tax (PAT), Sales, Market Capitalization and Interest coverage ratio (ICR) are taken from Capitaline Financial Database

Research tools used: The given analysis has been done with the help of SPSS .
Analysis Methodology

Discriminant analysis has been used for the purpose of our research. A typical two-variable discriminant equation takes the form of:

$$Y = a + k_1 x_1 + k_2 x_2$$

Where Y is the dependent variable and x_1 and x_2 are independent variables, k_1 and k_2 are the coefficients of the independent variables, and a is a constant.

Building a Model for Prediction/Classification

Assuming we have data on both the y and x variables of interest, we estimate coefficients of the model which is a linear equation of the form shown earlier, and use the coefficients to calculate the y value (discriminant score) – for any new data points that we want to classify into one of the groups. A decision rule is formulated for this process – to determine the cut off score, which is usually the midpoint of the mean discriminant scores of the two groups.

Accuracy of Classification:

To determine the accuracy of the model, classification of existing data points has been done using the equation. The outputs given by the classification matrix, tells us what percentage of the existing data points is correctly classified by this model.

This percentage is somewhat analogous to the R^2 in regression analysis (percentage of variation in dependent variable explained by the model). Of course, the actual predictive accuracy of the discriminant model may be less than the figure obtained by applying it to the data points on which it was based.

Analysis

The entire Research Analysis is divided into two proposed models, Model – 1 which predicts with extreme accuracy whether any company from the existing 50 companies in Nifty Index will be excluded or not in the near future and Model – 2 which predicts whether a particular company will be included as Nifty constituent or not in near future. These models are developed using discriminant analysis on past data on inclusion and exclusion of securities from Nifty Index.

Proposed Models are discriminant equations based on KPIs of a company as independent variables, KPIs are input for these equations. After plugging the value of KPIs in these equations, the value of dependent variable is calculated. The range of this value calculated is used to predict whether a particular

company will be excluded / not excluded from the Nifty index or included / not included in the Nifty Index.

Model is based on the following Key performance indicators (KPI) ie. Rate of growth of Profit after tax (ROG-PAT), Rate of growth of Net Sales (ROG-SALES), Rate of growth of Market Capitalization (ROG-MCAP), Interest Coverage Ratio (ICR) which are taken as independent variables in the discriminant equation.

PROPOSED MODEL-1 [EXCLUDED / NOT EXCLUDED]

Model–1 will predict whether any company from the existing 50 companies in Nifty Index will be excluded or not excluded in the near future. The proposed discriminant equation is

$Y = a1*(ROG\text{-}PAT - y1) + a2*(ROG\text{-}PAT - y2) + a3*(ROG\text{-}PAT\text{-} y3) + b1*(ROG\text{-}SALES - y1) + b2*(ROG\text{-}SALES - y2) + b3*(ROG\text{-}SALES - y3) + c1*(ROG\text{-}MCAP - y1) + c2*(ROG\text{-}MCAP\text{-} y2) + c3*(ROG\text{-}MCAP\text{-} y3) + d1*(ICR - y1) + d2*(ICR\text{-} y2) + d3*(ICR\text{-} y3) + E$

Where Rate of growth of Profit after tax (ROG-PAT), Rate of growth of Net Sales (ROG-SALES), Rate of growth of Market Capitalization (ROG-MCAP), Interest Coverage Ratio (ICR) are taken as independent variables, y1, y2, y3 denote the last three financial years. a1, a2, a3, b1, b2, b3, c1, c2, c3, d1, d2, d3 are coefficients of independent variables and E is a constant. Y is the dependent variable, the range of the value of calculated dependent variable is used to predict whether a particular company will be excluded / not excluded from the Nifty index

RESULTS

After executing the proposed discriminant equation on SPSS statistics software using the past data on exclusion of securities from Nifty index since its inception, we get the following results:

Coefficients of independent variables: Coefficients of independent variables a1, a2, a3, b1, b2, b3, c1, c2, c3, d1, d2, d3 and constant value E in the discriminant equation as calculated using SPSS statistics software are as shown in table1

Table 1

Canonical Discriminant Function Coefficients

	Function 1
GPATY 1	.001
GPATY 2	-.004
GPATY 3	-.001
GSALES Y 1	.002
GSALES Y 2	.003
GSALES Y 3	.007
ICRY 1	.000
ICRY 2	.001
ICRY 3	-.002
GMCAPY 1	.011
GMCAPY 2	.014
GMCAPY 3	.011
(Constant)	-.1.070

The final equation for model-1

The final equation for model-1 after plugging the coefficients and the constant value as calculated using SPSS statistics software is

Y=.001*(ROG-PAT - y1) -.004*(ROG-PAT - y2) - .001*(ROG-PAT- y3) + .002*(ROG-SALES - y1) + .003*(ROG-SALES - y2) +.007*(ROG-SALES - y3) +.011*(ROG-MCAP - y1)+.014*(ROG-MCAP- y2) + .011*(ROG-MCAP- y3) + .001*(ICR- y2) -.002*(ICR-y3) -1.070

Where Rate of growth of Profit after tax (ROG-PAT), Rate of growth of Net Sales (ROG-SALES), Rate of growth of Market Capitalization (ROG-MCAP), Interest Coverage Ratio (ICR) are taken as independent variables, y1, y2, y3 denote the last three financial years.

Prediction Criteria [Excluded / Not Excluded]

Table 2 : Functions at Group Centroids

STATUS	FUNCTION 1
1.00	-1.038
2.00	1.221

Unstandardized canconical discriminant functions evaluated at group means

Table 2 shows the prediction criteria

Y is the dependent variable which is calculated after plugging last three year financials in the final discriminant equation

Centre = (-1.038 + 1.221) / 2 = .0915, as calculated using functions at group centroids.

After plugging the last three years financials of any existing Nifty 50 constituent company in the equation,

If Y < .0915 than company may be excluded from NIFTY. If Y >.0915 then than company might not get excluded in the near future

Classification Results:

The Classification results show the percentage of Sample data that the proposed model shall be able to classify correctly.

Table 3 : Classification Results

		STATUS	Predicted Group Membership		
			1.00	2.00	Total
Original	Count	1.00	18	2	20
		2.00	2	15	17
		1.00	90.0	10.0	100.0
		2.00	11.8	88.2	100.0

a. 89.2% of original grouped cases correctly classified

Table 3 shows that this model can classify 89.2 % of the sample data correctly.

Statistical significance of the discriminant function

Table 4 : Wilks' Lambda

Test of Function(s)	Wilks' Lamba	Chi-square	df	Sig.
1	.428	24.640	12	.017

As seen in table 4 the WILKS'LAMBDA value comes out to be 0.428 which is indeed a low value and hence indicates better discriminating power of the model.

PROPOSED MODEL-2 [INCLUDED / NOT INCLUDED]

We again have developed another model: Model – 2 which predicts whether a particular company will be included as Nifty constituent or not in near future. The proposed discriminant equation for our second model is :

$$Y = a1*(ROG\text{-}PAT - y1) + a2*(ROG\text{-}PAT - y2) + a3*(ROG\text{-}PAT- y3) + b1*(ROG\text{-}SALES - y1) + b2*(ROG\text{-}SALES - y2) + b3*(ROG\text{-}SALES - y3) + c1*(ROG\text{-}MCAP - y1) + c2*(ROG\text{-}MCAP- y2) + c3*(ROG\text{-}MCAP- y3) + d1*(ICR - y1) + d2*(ICR- y2) + d3*(ICR- y3) + E$$

Where Rate of growth of Profit after tax (ROG-PAT), Rate of growth of Net Sales (ROG-SALES), Rate of growth of Market Capitalization (ROG-MCAP), Interest Coverage Ratio (ICR) are taken as independent

variables, y1, y2, y3 denote the last three financial years. a1, a2, a3, b1, b2, b3, c1, c2, c3, d1, d2, d3 are coefficients of independent variables and E is a constant. We then tried to predict Inclusion/Non inclusion on the basis of the range of the value of calculated dependent variable i.e..Y

RESULTS

After executing the proposed discriminant equation on SPSS statistics software using the past data on inclusion of securities from Nifty index since its inception, we got the following results:

Coefficients of independent variables
Coefficients of independent variables a1, a2, a3, b1, b2, b3, c1, c2, c3, d1, d2, d3 and constant value E in the discriminant equation as calculated using SPSS statistics software are as shown in table 5

Table 5

Canonical Discriminant Function Coefficients

	Function 1
GPATY 1	-.012
GPATY 2	.000
GPATY 3	.004
GSALES Y 1	-.015
GSALES Y 2	.031
GSALES Y 3	.025
ICRY 1	.001
ICRY 2	-.001
ICRY 3	.001
GMCAPY 1	.017
GMCAPY 2	.012
GMCAPY 3	-.002
(Constant)	-.1.994

Unstandardized coefficients

The final equation for model-2

The final equation for model-2 after plugging the coefficients and the constant value as calculated using SPSS statistics software is

$Y = .012*(ROG\text{-}PAT - y1) + .004*(ROG\text{-}PAT - y3) - .015*(ROG\text{-}SALES - y1) + .031*(ROG\text{-}SALES - y2) + .025*(ROG\text{-}SALES - y3) + .017*(ROG\text{-}MCAP - y1) + .012*(ROG\text{-}MCAP - y2) - .002*(ROG\text{-}MCAP - y3) + .001*(ICR - y1) - .001*(ICR - y2) + .001*(ICR - y3) - 1.994$

Where Rate of growth of Profit after tax (ROG-PAT), Rate of growth of Net Sales (ROG-SALES), Rate of growth of Market Capitalization (ROG-MCAP), Interest Coverage Ratio (ICR) are taken as independent variables, y1,

y2, y3 denote the last three financial years

Y is the dependent variable which is calculated after plugging last three year financials in the final discriminant equation

PREDICTION (INCLUDED / NOT INCLUDED)

Table 6   **Functions at Group Centroids**

STATUS	Function 1
1.00	-1.038
2.00	1.221

Unstandardized canconical discriminant functions evaluated at group means

Y is the dependent variable which is calculated after plugging last three financial years in the final discriminant equation

Centre = (1.864 + -1.598) / 2 = .113 as calculated using functions at group centroids

After plugging last financial data of three years of any particular company in the equation, If Y < .113 than company may be INCLUDED in NIFTY. If Y >.113 then than company might not get included

Percentage of Sample data that model could classify correctly

The Classification results in table 7 show the percentage of Sample data that the proposed model shall be able to classify correctly.

Table - 7 Classification Results*

		STATUS	Predicted Group Membership		
			1.00	2.00	Total
Original	Count	1.00	10	2	12
		2.00	0	14	14
	%	1.00	83.3	16.7	100.0
		2.00	.0	100.0	100.0

* 92.3% of original grouped cases correctly classified

Statistical significance of the discriminant function

Table - 8 Wilks' Lambda

Test of Function(s)	Wilks' Lamba	Chi-square	df	Sig.
1	.237	25.941	12	.011

As seen in table 8 the WILKS'LAMBDA value comes out to be 0.237

which is indeed a low value and hence indicates better discriminating power of the model.

Conclusion

The final outcomes of this research are as follows . Models that are built up on some key performance indicators of the securities such as rate of growth of Profit after tax (ROG-PAT), rate of growth of net Sales (ROG-SALES), rate of growth of Market Capitalization (ROG-MCAP) and Interest Coverage Ratio (ICR) can predict the effect of inclusion in and exclusion of securities from NIFTY index. The models use the Discriminant analysis on the SPSS Platform.

Eligibility criteria for inclusion and exclusion of securities in NIFTY index are Impact cost (Liquidity), Market Capitalization, Domicile, Eligible types of Securities and free float. Share prices tend to increase prior to inclusion and decrease prior to exclusion from the Index.

A long (buy) position if taken in the securities that are predicted by the model to be included in the index will help in generating higher returns as its financial performance will improve before the inclusion leading to higher stock price as well. Alternately, a short (sell) position taken in the securities that are predicted by the model to be excluded from the index will also be a good bet as its financial performance will degrade before the exclusion leading to lower stock price.

References

Agarwal, S., J. Driscoll, X. Gabaix, and D. Laibson, (2007), *The age of reason: Financial Decisions over the Lifecycle, mimeo*, Harvard University.

Agnew, J., and L. Szykman, (2005), Asset allocation and information overload: The influence of information display, asset choice and investor experience, J*ournal of Behavioral Finance*,6

Campbell, J., A. Lo and A.C. Mackinlay, (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Cocco, J., F. Gomes, and P. Maenhout, (2005), Consumption and Portfolio Choice over the lifecycle, *Review of Financial Studies,* 18.

Dimson, Elroy, Paul Marsh, and Mike Staunton. (2008). "The Worldwide Equity Premium: A Smaller Puzzle." In *Handbook of the Equity Risk Premium, ed. Rajnish Mehra*,. Amsterdam: Elsevier.

Guiso, L., P. Sapienza, and L. Zingales, (2005), Trusting the stock market, *NBER Working Paper* , 11648.

Guiso, L., and T. Jappelli, (2005), Awareness and stock market participation, *Review of Finance*, 9.

Rietz, Thomas A. (1988). "The Equity Risk Premium: A Solution." *Journal of Monetary Economics.*

Mehra, Rajnish, and Edward C. Prescott. (1985). "The Equity Premium: A Puzzle." *Journal of Monetary Economics,*

Haliassos, M., and C. Bertaut, (1995), Why do so few hold stocks?, *Economic Journal,* 105.

Hicks, J.R. (1939). *Value and capital.* Oxford University Press: London.

Leuthold, R.M. (1983). Commercial use and speculative measures of the livestock commodity futures markets. *Journal of Futures Markets* 3, 113-135.

Thorbecke, W., June (1997) "On Stock Markets Returns and Monetary Policy", *Journal of Finance*, Vol. 76.