IIMB-WP N0. 596/2019

WORKING PAPER NO: 596

Predictability of Equity Risk Premium in Indian Equity Markets

V. Ravi Anshuman
Professor
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
anshuman@iimb.ac.in

Rajdeep Sharma
Doctoral Student
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
rajdeep.sharma18@iimb.ac.in

Prateek Jain
Doctoral Student
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
jain.prateek18@iimb.ac.in

Arnab Biswas
Doctoral Student
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
arnab.biswas18@iimb.ac.in

Year of Publication – September 2019

Predictability of Equity Risk Premium in Indian Equity Markets

Abstract

We show that the historical mean of the equity risk premium is consistently a more accurate out-of-sample predictor of future equity risk premium in Indian equity markets. Under certain variations of the in-sample period length, dividend payout and the mean combination forecast have better predictive power than the historical mean equity risk premium. Finally, we find that predictions based on more recent information are, on average, more accurate than those based on the entire history of observations. We estimate that the (geometric) average annual equity risk premium of NIFTY 500 index for the period June 2000 to March 2018 is 7.78%[1].

JEL Classification: G12, G14, C22

Keywords: equity risk premium, stock returns, forecasting, predicting returns, Indian market index, NIFTY, asset pricing

---

[1] Table 1 in appendix reports summary statistics for NIFTY 500 index returns, various measures of risk-free rate and BBB-rated 10-year corporate bonds (lowest investment grade). The returns and yields are monthly, and not annualized.
Table 2 in appendix reports the (arithmetic, geometric and continuously compounded) annualized average returns for the NIFTY 500 Total Returns Index in Panel A and arithmetic and geometric ERP estimates using various risk-free proxies in Panel B.

1. Introduction

The equity risk premium (ERP) of the market portfolio is widely used in the field of corporate finance, valuation, and portfolio management. Not surprisingly, predictability of the equity risk premium is of great importance to researchers. Researchers have used different economic variables as potential predictors of the ERP. Welch and Goyal (2008) conduct a comprehensive analysis of ERP prediction models for the U.S., based on predictors such as earnings to price ratio, dividend to price ratio, term spread, book to market ratio etc.[2]. They find that most of the predictors of ERP have reasonable in-sample predictive power but suffer from weak out-of-sample predictive power, relative to a benchmark predictor based on the historical mean of ERP. In general, all predictors, including the historical mean ERP, are poor predictors of ERP; it is just the case that, on average, the historical mean ERP's out-of-sample predictive performance is less inferior to that of other predictors. Hence, on a relative basis, the historical mean ERP is still acknowledged to be the best predictor of the future ERP[3].

In the Indian context, Narayan and Bannigidadmath (2015) analyze the predictive power of the standard predictors of the equity risk premium for a set of stylized portfolios sorted on the basis of industry, size, and value. They find evidence of industry return predictability during expansions and evidence of book-to-market and size portfolio return predictability during recession.

[2] Polk, Thompson and Vuolteenaho (2006) forecast cross-sectional equity risk premium and find that earnings yield explains a large fraction of the time series variation in ERP. Their econometric analysis is, however, on an individual stock level rather than on a diversified portfolio.
[3] In a novel approach, Rapach, Strauss, and Zhou (2010) determine a weighted average of the ERP predictions of each predictor variable. In other words, the traditionally accepted benchmark of the historical mean ERP can potentially be improved upon by using a weighted average ERP prediction.

While our study also deals with similar issues as the study by Narayanan and Bannigidadmath (2015), our study is distinct in two significant ways. First, we focus our analysis on the market index as compared to the stylized portfolios examined in their study; thus, our study has greater applicability from a capital budgeting perspective. Second, we use data from the period 2000-2018, which reflects a more mature phase of the liberalization of the Indian economy; thus, our study provides a more contemporary perspective and is free from the effects of significant structural shifts in the Indian economy that occurred during the first phase of liberalization (1991-2001).

In this study, we estimate the predictability of ERP for the NIFTY 500 Total Returns Index (TRI), which is a value-weighted index and proxies for the market portfolio in Indian equity markets. The return on the market portfolio (NIFTY 500) would be the return that investors would realize if they had invested in the "average risk"[4] equity investment. NIFTY 500 represents the top 500 companies based on full market capitalization and 94% of total free-float market capitalization[5] of the stocks listed on NSE as on March 31, 2016. It is, therefore, a good choice for the diversified market portfolio. The independent variables that we employ in our study to explain variation in ERP are indicators of value (book to market, earnings to price), cash flow generating capacity (dividend yield, dividend price and dividend

[4] "Average risk" means the value-weighted risk of the constituent 500 stocks of NIFTY 500.
[5] Source: NSE Indices website (https://www.niftyindices.com/indices/equity/broad-based-indices/nifty-500)
NIFTY 500 represents about 94% of the free float market capitalization of the stocks listed on NSE as on March 31, 2016.

payout), volatility (variance of daily returns) and two macroeconomic indicators, long-term inflation (term spread) and credit default risk (default spread).

We benchmark our prediction of the ERP against the historical mean of the ERP realized over the in-sample window (historical mean ERP). This helps us evaluate the performance of predictor variables vis-à-vis the historical mean ERP benchmark. Essentially, we test whether investors can predict ERP based on the information contained in these economic variables or simply use the trend of returns (historical mean ERP) to predict ERP. We also consider the average of the predicted values from each predictor variable, as another predictor variable (the mean combination forecast).

We consider two different approaches for selecting the in-sample period for the regressions. In the first method, called rolling regression method, the in-sample period is of fixed length of latest 48 months. In the second method, called augmented regression method, we start with an initial in-sample period covering of length 48 months and for each subsequent future period, the in-sample period length increases by one month.

Our results show that, when using the rolling regression method with in-sample period length of 48 months, dividend payout outperforms historical mean ERP in out-of-sample forecasting. Upon varying the length of in-sample periods, we observe that divided yield and mean combination forecast have superior predictive ability when compared to historical mean ERP. However, when we use the augmented regression method, all predictor variables are less accurate than historical mean ERP. The out-of-sample forecasting errors for our model increase significantly during the recession period from January 2008 to June 2009.

We use different measures to quantify the forecasting accuracy of the predictor variables for both rolling and augmented regression methods. Upon ranking the predictor variables based on these measures, we find that the mean combination forecast, although less accurate than predictions based on historical mean ERP, is a relatively stable out-of-sample predictor.

Our study is useful for all industry practitioners who require estimates of the market risk premium (MRP), e.g., for regulators computing the cost of capital in a regulated industry or portfolio managers wishing to allocate their client's wealth across a portfolio of the risk-free asset and the market portfolio. It is also informative about the debate on the advantages of investing social security funds in stocks (Fama and French, 2002). In addition, foreign fund flows to India have been steadily increasing over the last decade as the Indian equity markets offer an attractive alternative to foreign investors. Thus, the estimation of ERP is of great interest to wide spectrum of market practitioners. The models we develop in this paper are tractable and can be used for capital budgeting, relative valuation, and portfolio management disciplines.

The rest of the paper is organized as follows. Section 2 describes the data. We discuss the empirical methodology in Section 3. In Section 4, we discuss the empirical results and perform sensitivity analysis of rolling regression method to varying in-sample period length. Section 5 concludes our study.

2. Data and Variables

We obtain the dataset from Bloomberg. Following is the description of our dataset, transformations done on raw data and the time period of available data:

Monthly Returns (Rm): The natural logarithm of the ratio of the total returns index level on the last trading day of the current month divided by the total returns index level of the last trading day of the previous month.

Risk-Free Rate (Rf): Rf is the rate of return of an investment with zero risk. The annualized risk-free rate for any month is computed by taking the arithmetic mean of the daily yields of 30-day Treasury bill in the month. The conversion to a monthly rate is done by multiplying the annualized rate for the month with the actual number of days in that month and dividing by 365. This monthly rate is used as the risk-free rate.

Equity Risk Premium (ERP): Equity Risk Premium is the difference between the Monthly Returns and the continuously compounded risk-free rate, i.e., log (1+ Rf**)**. The monthly ERP is computed for all months starting from June 2000 till March 2018.

Term Spread (TERM_SPR): Term Spread is the difference between yield to maturities of long-term bonds and short-term bonds. This measure encapsulates the outlook on long term inflation in the economy. Long-term bond is the 10-year AAA Government Bond and short-term bond is the 91-day Treasury Bill. YTM for both long-term bonds and T-bills are annualized rates, observed on each trading day. The YTM of long-term bonds and T-bills for any month is computed by taking the average of their YTMs observed on each trading day of that month. The difference between the monthly YTMs of long-term bonds and T-bills is the observed term spread for that month. Term spread is computed for all months starting from June 2000 till March 2018.

Default Spread (DEF_SPR): Default Spread captures the YTM difference between BBB-rated l0-year Corporate bond and AAA-rated 10-year Government bonds. This measure encapsulates the outlook on default risk in the corporate sector of the economy. YTM for both corporate and government bonds are annualized rates, observed on each trading day. The annualized YTM of BBB-rated l0-year Corporate bond and AAA-rated 10-year Government bond for any month is computed by taking the average of their YTMs observed on each trading day of that month. The difference between the YTM of BBB-rated l0-year Corporate bond and AAA-rated 10-year Government bond, for any month, is the expected default spread for that month. The default spread is computed for all months starting from July 2004 till March 2018.

AAA is the highest rated bond and BBB is the lowest investment grade (White, 2010). Hence, this difference will capture the spread across the investment spectrum, which is expected to rise in the times of financial crises and narrow down when the economy is in good shape.

Earnings: Earnings is the market capitalization weighted trailing-twelve-month (TTM) earnings of each of the 500 constituent stocks of the NIFTY 500 index. TTM earnings are calculated as of the last trading day of each month.

Earnings to Price (E/P): Earnings to Price for a month is the natural logarithm of the ratio of earnings to the Price level of the NIFTY 500 index[6]. Price is the index level of the NIFTY 500 index on the last trading day for the same month. From the data available, we calculate earnings to price ratio for all months starting from June 2002 to March 2018.

Book to Market (BOOK_MKT): Book to Market is the ratio of book value to the Price level of the NIFTY 500 index. Book Value is the market capitalization weighted book value of each of the 500 constituent stocks of the NIFTY 500 index. For months from January to September in any year, we consider book value as of 31st March of the previous calendar year for the ratio computation. For months from October till December of any year, we consider book Value as of 31st March of the same calendar year for the ratio computation. Price is the index level of the NIFTY 500 index on the last trading day for the same month corresponding to which compute the book value. We calculate the book to market for all months starting from October 2003 till March 2018.

Dividend Price (DIV_PRICE): Dividend Price is the natural logarithm of the ratio of dividends to price of the NIFTY 500 index. Dividends are the market-capitalization weighted TTM dividends of each of the 500 constituent stocks of the NIFTY 500 index. TTM dividends are calculated as of the last trading day of each month. Price is the index level of the NIFTY 500 index on the last trading day of

[6] NIFTY 500 index is the level of the NIFTY 500 Total Return Index net of dividends, interest, rights offerings and other distributions.

the same month corresponding to which we compute the TTM dividends. We calculate the dividend yield for all months starting from June 2002 till March 2018.

Dividend Yield (DIV_YIELD): Dividend Yield is the natural logarithm of the ratio of dividends to price of the NIFTY 500 index calculated on the last trading day of each month. Dividends are the market-capitalization weighted TTM dividends of each of the 500 constituent stocks of the NIFTY 500 index. TTM dividends are calculated as of the last trading day of each month. For each of the months in which we calculate TTM dividends, we also compute the twelve-month lagged price level of the NIFTY 500 index. This price level is then used to compute dividend yield. We calculate dividend price for all months starting from June 2002 till March 2018.

Dividend Payout (DIV_PAYOUT): Dividend Payout is the natural logarithm of the ratio of dividends to earnings of the NIFTY 500 index calculated on the last trading day of each month. Dividends are the market-capitalization weighted TTM dividends of each of the 500 constituent stocks of the NIFTY 500 index. TTM dividends are calculated as of the last trading day of each month. Earnings are the market -capitalization weighted TTM earnings of each of the 500 constituent stocks of the NIFTY 500 index. TTM earnings are calculated as of the last trading day for the same month corresponding to which we compute the TTM dividends. We calculate dividend payout for all months starting from June 2002 till March 2018.

Stock Variance (SVAR): Stock Variance is the measure of daily volatility in returns of the NIFTY 500 index. Variance in the daily log returns of the NIFTY 500 index is calculated for every month in our data set. We calculate the stock variance for all months starting from June 2000 till March 2018.

Historical Mean ERP (HIST_MEAN): Historical Mean ERP is the average of realized ERPs of past months, used as a benchmark for comparing prediction accuracy of predictor variables. In rolling regression method, historical mean ERP is the mean of realized ERP of latest 48 months, whereas, in augmented regression, it is the mean of all monthly realized ERPs from the starting month of the ERP time series till the forecasting period.

3. Methodology

Using the Augmented Dicky Fuller (ADF) test, we test all the variables for stationarity. The p-values of the ADF unit root test are reported in Table 3. We observe that dividend price, dividend yield, dividend payout, term spread, and default spread are non-stationary. The p-value of the test statistic for these variables is greater than 0.05, hence we fail to reject the null hypothesis of non-stationarity for these variables at a significance level of 0.05. In order to get reliable test statistics for the regressions, we take the first difference of non-stationary variables before using them in our regression model. The stationary predictor variables are used in the level form in our regressions.

[Insert Table 3 here]

3.1 Regression Methods

We use two different methods of selecting the in-sample period for training the regression model and obtaining the ARIMAX estimates, namely the rolling regression method and the augmented regression method.

3.1.1 Rolling Regression Method

In this method, we assess the constancy of the model's forecasting accuracy over a rolling window, keeping the in-sample size fixed at 48 months, while rolling through the entire dataset. The starting month of the in-sample period of the first regression for different predictor variables may vary due to constraints on their data availability. E.g., default spread data is available only from July 2004 onwards, hence the first training window ranges from August[7] 2004 to July 2008 (48 months). The in-sample period is then rolled forward by one month till we reach the end of our dataset (i.e., March 2018).

For each in-sample period, we compute the autocorrelation order of ERP (the dependent variable) and the predictor variable. We find that the autocorrelation order for ERP is zero for 98% of the rolling in-sample periods. For the predictor variables, we find that the autocorrelation order varies for different rolling in-sample periods. However, none of the autocorrelation orders, for any predictor variable, is greater than 4.

Based on these statistical properties of our dataset, we develop the following ARIMAX model that fits our in-sample period data.

For non-stationary predictor variables,

$$ERP_t = \beta_0 + \beta_1 \times \Delta Predictor_{t-1} + \beta_2 \times \Delta Predictor_{t-2} + \ldots.. + \beta_{\mathrm{p}} \times \Delta Predictor_{t-\mathrm{k}} + u_t \quad (1)$$

where u_t follows an ARMA (p,q) process

[7] We lose one month's data due to first differencing to correct for non-stationarity.

For stationary predictor variables,

$$ERP_t = \beta_0 + \beta_1 \times Predictor_{t-1} + \beta_2 \times Predictor_{t-2} + \ldots\ldots + \beta_{\mathrm{p}} \times Predictor_{t-\mathrm{k}} + u_t \qquad (2)$$

where u_t follows an ARMA (p,q) process,

and k = 1, 2, 3, 4, 5; we choose the best model out of these based on the minimum AICc criterion[8].

For each in-sample period, we test for heteroskedasticity in monthly ERP by using the Ljung-Box test on the in-sample squared residuals, up to a lag order of 5, obtained from the models (1) or (2). We find no evidence of heteroskedasticity in any of the in-sample periods in rolling regression method and therefore we do not estimate a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model for conditional volatility.

We use the ARIMAX model in our regressions of ERP on each of the predictor variables using regression models (1) or (2) to estimate the coefficients using the method of maximum likelihood. We then use the regression coefficients to obtain the one step ahead forecast from the estimated model. For instance, if the first stationary observation of a predictor variable starts at time period (t+1), then ARIMAX regressions are run for months (t+6, t+53)[*], (t+7, t+54), (t+8, t+55) till

---

[8] For small sample sizes, there is a significant probability that AIC criterion will pick models that have too many parameters, i.e. that there will be overfitting. To avoid overfitting, we use AICc which is AIC with a correction for small sample sizes. AICc has a penalty term for the number of parameters.

$AICc = AIC + \frac{2k^2+2k}{n-k-1}$, where n is the sample size and k denotes the number of parameters.

[*] Since we test for a maximum of 5 lags of predictor variable in our ARIMAX model and choose the one with minimum AICC criterion, hence, corresponding to each predictor variable, we lose 5 initial observations of ERP, i.e., (t+1) to (t+5) in this case.

we reach the end point in our dataset (i.e. March 2018). For each of these regressions, we estimate the out-of-sample ERP for the month immediately following the last month of the in-sample period of the regression. For example, if the in-sample period for obtaining the ARIMAX estimates ranges from month (t+6) to (t+53), we predict the ERP (using the estimates obtained) for the month t+54.

For rolling regression method, we compute the historical mean ERP by taking an arithmetic average of the realized ERPs over the in-sample period, i.e., the latest 48 months. The mean thus obtained is the subsequent month's historical mean ERP estimate E.g., the historical mean ERP estimate for month (t+54) is the arithmetic average of realized ERPs of months (t+6) to (t+53). Hence, we compute the historical mean ERP on a rolling basis and use these values as predicted ERPs (using HIST_MEAN as predictor variable) for subsequent months.

3.1.2 Augmented Regression Method

In the augmented regression method, the starting month of the in-sample periods of all regressions remain fixed, and the in-sample period length increases by one month as we advance into the next month.

For each predictor variable, we start from the month of its first observation, with an in-sample period length of 48 months for the first regression. The starting month of the subsequent in-sample periods is fixed (same as the first in-sample period) but may vary across different predictor variables due to constraints on their data availability, e.g., default spread data is available only from July 2004 onwards, hence the first training window ranges from August 2004 to July 2008 (48 months), losing one observation due to first differencing to remove non-stationarity.

For each in-sample period, we compute the autocorrelation order of ERP (the dependent variable) and the predictor variable. We find that the autocorrelation order for ERP is zero for 74% of our in-sample periods in the augmented regression method. For the predictor variables, we find that the autocorrelation order varies with different in-sample periods in the augmented regression method. However, none of the autocorrelation orders, for any predictor variable, is greater than 4. Hence, in the augmented regression method, we employ the same set of ARIMAX models (1) and (2), which we use in rolling regression method, where the lag order of the model is based on the minimum value of the AICc criteria.

For each in-sample period, we test for heteroskedasticity in monthly ERP by using the Ljung-Box test on the in-sample squared residuals, up to a lag order of 5, obtained from the models (1) or (2). We find no evidence of heteroskedasticity for more than 95 percent of the in-sample periods across the predictor variables and therefore we do not estimate a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model for conditional volatility.

We use the ARIMAX model in our regressions of ERP on each of the predictor variables using regression models (1) or (2) to estimate the coefficients using the method of maximum likelihood. We then use the regression coefficients to obtain the one step ahead forecast from the estimated model. For instance, if the first stationary observation of a predictor variable starts at time period (t+1), then ARIMAX regressions are run for months (t+6, t+53), (t+6, t+54), (t+6, t+55) till we reach the end point in our dataset (i.e., March 2018). For each of these regressions, we estimate the out-of-sample ERP for the month immediately following the last month of the in-sample period of the regression. For example, if the in-sample period for obtaining the ARIMAX estimates ranges from month (t+6) to (t+95), we predict the ERP (using the estimates obtained) for the month t+96.

For augmented regression method, we compute the historical mean ERP by taking an arithmetic average of the realized ERPs of the entire dataset, till the forecasting period. The mean thus obtained is the subsequent month's historical mean ERP estimate, e.g., the historical mean ERP estimate for month (t+96) is the arithmetic average of realized ERPs of months (t+6) to (t+95). We use these values as predicted ERPs (using HIST_MEAN as predictor variable) for subsequent months. In the augmented regression method, we compute the historical mean ERP over a period of increasing length, whereas in rolling regression method, the averaging period length is kept fixed (latest 48 months). The number of regressions for any predictor variable is same for both rolling and augmented regression methods. The number of regressions for each predictor variable are given in Table 4.

[Insert Table 4 here]

For each month in which ERP prediction(s) are obtained from univariate regressions (corresponding to each of the predictor variables), we compute an arithmetic average of these ERP predictions. Henceforth, we will refer to this average as the mean combination forecast (MEAN_COMB). The average of these ERP predictions in a month is defined as mean combination forecast for that month. We aim to pick up economically meaningful changes from all the eight economic variables through this average forecast and significantly improve the out-of-sample predictive performance relative to individual predictor variables (Narayan and Bannigidadmath, 2015). In other words, the mean combination forecast measure might result in lower prediction error (absolute value of residuals), relative to the individual predictors. We get mean combination forecast under both, the rolling and the augmented regression methods.

3.2 Out-of-Sample Statistics

Consistent with our motivation in this paper, we compare the out-of-sample predictive power of the predictor variables and mean combination forecast with historical mean ERP. We analyze the results of both, rolling and augmented regression methods. The measures used to rank the predictor variables, based on their predictive power, are mentioned below. These measures have been used in earlier studies on equity risk premium (Goyal and Welch, 2008).

3.2.1 Difference of Root Mean Square Errors (∆RMSE)

We compute the Root Mean Square Error (RMSE) of the predictions of the individual predictor variables, mean combination forecast and historical mean ERP. Root Mean Square Error (RMSE) is the square root of the arithmetic average of the square of the prediction errors, where prediction error is simply the difference between out-of-sample predicted ERP and the observed ERP for that month. ∆RMSE for each of the predictor variables and mean combination forecast is defined as per equation (3).

$$\Delta RMSE_{Predictor} = (RMSE_{Historical}) - (RMSE_{Predictor}) \quad (3)$$

where, $RMSE_{Historical}$ denotes RMSE of the ERP predictions using historical mean ERP and $RMSE_{Predictor}$ denotes RMSE of the ERP predictions using predictor variables and mean combination forecast.

3.2.2 R^2

R^2 is defined as one minus the ratio of mean square error of predicted ERPs using predictor variables and mean combination forecast to the mean square error of ERP prediction using historical mean ERP. The R^2 for each of the predictor variables and mean combination forecast is obtained as per equation (4).

$$R^2{}_{Predictor} = 1 - \frac{MSE_{Predictor}}{MSE_{Historical}} \quad (4)$$

where, $MSE_{Historical}$ denotes the mean squared error of predictions using the historical mean ERPs and the $MSE_{Predictor}$ denotes mean squared error of predictions using the predictor variables and mean combination forecast.

Both the $\Delta RMSE$ and R^2 are relative measures of predictive performance as they are benchmarked on the performance of historical mean ERP. For historical mean ERP, both these measures are equal to zero by constriction, as can be verified by equations (3) and (4).

3.2.3 Mean Absolute Percentage Error (MAPE)

MAPE is the mean absolute percentage error of our predictions with respect to the actual ERPs for the out-of-sample period. We compute MAPE for each of the predictor variables, mean combination forecast and historical mean ERP. MAPE is calculated as per equation (5),

$$MAPE(\%) = \frac{1}{N} \times \frac{|Predicted\ ERP - Realized\ ERP|}{|Realized\ ERP|} \times 100 \qquad (5)$$

where N is the number of out-of-sample ERP predictions.

3.2.4 Percent Sign Correctly Predicted (PSCP)

For each predictor variable, PSCP counts the number of out-of-sample predictions which have the same sign as their corresponding realized ERPs. We compute PSCP for each of the predictor variables, mean combination forecast and historical mean ERP. PSCP is calculated as per equation (6)

$$PCSP(\%) = \frac{1}{N} \times (Number\ of\ correct\ sign\ predictions) \times 100 \qquad (6)$$

where N is the number of out-of-sample ERP predictions.

To complement MAPE, which measures the accuracy in terms of absolute percentage error, we check for the directional accuracy of the ERP predictions through PSCP.

3.3 Ranking Framework for Predictor Variables

We compare the predictive power of the predictor variables, mean combination forecast and historical mean ERP for both rolling and augmented regression methods. The out-of-sample predictive power of the predictor variables, mean combination forecast, and historical mean ERP is compared based on ΔRMSE, R^2, MAPE and PSCP.

ΔRMSE and R^2 are measures of the error in prediction when compared to the prediction based on historical mean ERP. Positive and greater value of ΔRMSE or R^2 implies that the predictor variable or mean combination forecast is more accurate than historical mean ERP in out-of-sample forecasting. Negative and lower values of ΔRMSE or and R^2 imply the opposite. Hence, predictor variables having positive and higher values of ΔRMSE and R^2 are ranked higher (lower numerical value of rank). ΔRMSE and R^2 are similar measures of ranking as both variables are functions of mean square error (MSE) of the predictor variable and historical mean ERP. However, a predictor variable may be ranked differently under both these measures.

We refer to ΔRMSE and R^2 as measures of relative ranking as the values of these measures are relative to the historical mean ERP predictions. For historical mean ERP, both these measures are equal to zero by construction.

Since MAPE measures the mean absolute percentage error, lower values of MAPE indicate a lesser deviation from realized ERPs. However, we cannot use MAPE in isolation to rank the variables based on predictive ability. This is because although the absolute deviation in prediction of ERP may be small, but the sign prediction may be opposite to that of the realized ERP. Taking only the absolute percentage deviation masks the directional accuracy of predictions. To complement MAPE, we use PSCP.

PSCP is a measure of the percentage of times the predicted value of ERP is of the same sign as that of the realized ERP. This lets us know the precision of the sign predictions. Consequently, variables with low MAPE and high PSCP values are ranked higher (lower numerical value of rank) than variables with high MAPE and low PSCP respectively.

MAPE and PSCP are absolute measures of predictive performance as, contrary to ΔRMSE and R^2, they do not require to be benchmarked against the MAPE and PSCP values of historical mean ERP. Hence, for historical mean ERP, both these measures are not equal to zero by construction.

We divide the four measures of predictive performance into two groups. First group (henceforth, Relative Group), consists of the relative ranking measures, ΔRMSE and R^2 and second group (henceforth, Absolute Group), consists of the absolute ranking measures, MAPE and PSCP.

We comment on the accuracy of predictions by computing the ranks of the predictor variables under each of the four measures. Therefore, for each predictor variable, we have four ranks corresponding to each of the measures (two each under Relative and Absolute Groups). Finally, for both Relative and Absolute Groups, we compute a composite rank for each variable by ranking them based on average of the ranks obtained under the two measures within each group.

In the results section, we capture the forecasting performance of the predictor variables, mean combination forecast and the historical mean ERP under both rolling and augmented regression methods. Finally, we compare the predictive performance of the two regression methods.

4. Results

We can summarize our empirical results by three main findings. First, historical mean of the equity risk premium is consistently a more accurate out-of-sample predictor of future equity risk premium in Indian equity markets. Second, except for dividend payout, none of the predictor variables or mean combination forecast could predict out-of-sample ERP more accurately than the historical mean ERP. Dividend payout performs better than historical mean ERP under rolling regression method (only Relative Group). It is inferior to historical mean ERP under augmented regression method. Finally, predictions obtained from the rolling regression method are, on average, more accurate than the augmented regression method for most of the predictor variables.

We discuss the detailed results under rolling regression method first, followed by the augmented regression method. Subsequently, we compare the relative accuracy of out-of-sample predictions of each predictor variable under both regression methods.

4.1 Rolling Regression

When we compare the performance of predictor variables under Relative Group, we see that except dividend payout, all the predictor variables have lesser forecasting accuracy than historical mean ERP. This is evidenced by the positive value of both ΔRMSE and R^2 for dividend payout and negative values for all the other predictor variables. There is a little difference in individual rankings of the variables under both these measures. Among the predictor variables, dividend payout is the most accurate while term spread is the least accurate predictor, based on composite rank (average rank of both measures). Historical mean ERP is second to dividend payout in composite rankings (Table 5 in appendix).

[Insert Table 5 here]

The performance of the predictor variables under Absolute Group is reported in Table 6 in appendix. Since MAPE and PSCP measures complement each other, we get a true sense of the performance of predictor variables under Absolute Group by looking at the composite rank. Mean combination forecast is the highest ranked variable, followed by dividend payout and historical mean ERP (both rank second). It is important to note here that although dividend payout performs slightly worse than mean combination forecast under Absolute Group, its performance is better than both, the historical mean ERP and the mean combination forecast, under Relative Group.

[Insert Table 6 here]

4.1.1 Sensitivity of Results Under Rolling Regression Method to varying in-sample period lengths

We test the robustness of the results obtained under rolling regression method by using three different in-sample period lengths of 36, 48 and 60 months respectively. We find that the forecasting accuracy of the predictor variables varies with respect to in-sample period lengths. For the 36-month in-sample period, based on ΔRMSE, mean combination forecast (rank = 1) performs better than historical mean ERP (rank = 2) while dividend payout is ranked third. For the 48-month in-sample period, dividend payout is the highest ranked variable while historical mean ERP is ranked second, followed by mean combination forecast. Further, for the 60-month in-sample period, none of the predictor variables are more accurate than the historical mean ERP. Mean combination forecast is ranked fourth while dividend payout is ranked eighth.

Therefore, the results of the rolling regression method are sensitive to the choice of in-sample period length. Historical mean ERP's performance is consistent across different in-sample period lengths. The rankings of the predictors based on ΔRMSE, across different choices of in-sample period length are given Table 7 in appendix.

[Insert Table 7 here]

4.2 Augmented Regression

When we compare the performance of predictor variables under Relative Group, historical mean ERP is the most accurate predictor (Table 8 in appendix). The relative performance of the predictor variables changes under this regression method, when compared with rolling regression method. We find that, among the predictor variables, historical mean ERP is the most accurate predictor followed by default spread, whereas stock variance is the least accurate predictor based on composite rank. Dividend payout is ranked seventh based on composite rank, compared to rank one in Relative Group under rolling regression method. Based on composite rank, mean combination forecast is ranked third under both augmented and rolling regression methods in Relative Group.

[Insert Table 8 here]

The composite ranking of the predictor variables under Absolute Group using the augmented regression method is reported in Table 9 in appendix. We find that mean combination forecast is the highest ranked variable, followed by the historical mean ERP and stock variance (both ranked second). Default spread, which is ranked second under Relative Group, performs poorly in Absolute Group (rank = 7).

[Insert Table 9 here]

4.3 Comparison of Augmented and Rolling Regressions

We compare the ranking variation in the predictor variables across both the regression methods and ranking measures under Absolute and Relative Groups. We compute two measures that capture the average rank of the predictor variable (Mean Rank) and the variation in ranks (Variance in Ranks). Mean Rank is calculated by re-ranking the simple average of the composite ranks across regression methods and Absolute and Relative Groups. Variance in Ranks, as the name suggests is the sample variance of the ranks across regression methods and Absolute and Relative Groups.

We find that book to market has the lowest variation in ranks (along with historical mean ERP) but given its low mean rank (rank=6), it is a consistently poor predictor of ERP. Historical mean ERP also has the lowest variation in ranks but because of its highest mean rank (rank = 1), it is the most consistent predictor of ERP. Closely following the historical mean ERP is the mean combination forecast, which has the second lowest variation in ranks and is second in mean rank (Table 10 in appendix). Hence, the mean combination is a highly stable predictor of ERP. Dividend payout's mean rank is three, which is due to its inferior predictive performance under augmented regression method as compared to rolling regression method. All the other predictor variables fail to predict as accurately and consistently as historical mean across both the regression methods and Absolute and Relative Groups.

[Insert Table 10 here]

In order to compare the performance of predictor variables under augmented and rolling regression methods, we compute the difference (augmented minus rolling) of the MAPE and PSCP values, for each of the predictor variables, under the two regression methods (Table 11). For any predictor variable, a positive value of MAPE difference and negative value of PSCP difference indicates that the predictor variable is more accurate in predicting ERP under rolling regression method. We find that for both measures, variables dividend payout and earnings to price ratio have better performance in rolling regression than augmented regression. This is because the signs on the differences under MAPE and PSCP measures are alternating in nature. Hence, dividend payout, which outperforms the historical mean ERP under rolling regression method, also has superior predictive power under the rolling regression method than the augmented regression method. Also, we find that most predictors, on average, are more accurate at forecasting ERP under the rolling regression method as evidenced by the sign of the average values. Hence, accuracy of predictions based on more recent information are, on average, higher.

[Insert Table 11 here]

Finally, the time period during which we tracked the out-of-sample predictive performance of the predictor variables also includes the 2008-09 recession period. Indian equity markets were impacted adversely by the crisis and our econometric model has poor predictive power during this period as compared to the other periods. The mean squared errors of predictions increase significantly during the recession period from January 2008 to June 2009 and decline to the pre-crisis levels by late 2009.

5. Conclusion

In this paper, we investigate the predictability of returns of the market portfolio (NIFTY 500) by employing numerous economic variables that had been suggested by literature to have reasonable predictive accuracy. We test the out-of-sample predictive power of these variables in the Indian equity markets from the period November 2004 till March 2018 by using ARIMAX models of the economic variables on the ERP. We also compare the predictive power of these variables under two different methods of regression. The rolling regression uses only the recent market information for ERP prediction while the augmented regression uses all the market information available till the forecast period.

We propose an innovative ranking framework for predictor variables based on their forecasting performance by using various performance measures. Based on these measures, we also conclude that, on average, rolling regression method has superior performance than the augmented regression method.

Under the rolling regression method, we find that dividend payout outperforms historical mean ERP in the out-of-sample period. However, the performance of predictor variables is sensitive to the choice of in-sample period length in the rolling regression method. Under the augmented regression method, we find that all the predictor variables perform worse than the historical mean ERP. We find that, on average, the prediction accuracy of predictor variables is higher under the rolling regression method as compared to the augmented regression method. Hence, more recent market information plays a greater role than the entire past information. Compared to our predictor variables, we find that historical mean equity risk premium is consistently more accurate out-of-sample predictor of equity risk premium across both regression methods.

Our study has important implications for industry practitioners in the realm of portfolio management and for regulators and treasury managers interested in estimating the cost of capital over a given horizon. The model that we develop in this paper is tractable and easy to implement. An investor can also look at a weighted combination of predictions made by each of the predictor variables used in our study to see if it has higher accuracy in predicting ERP than dividend payout, simple mean combination forecast or the historical mean ERP. However, finding the optimal set of dynamic (time-varying) weights could be challenging computationally and is a promising scope for further research.

6. References

1. Chong, J., & Hurn, A. S. (n.d.). Testing for Speculative Bubbles: Revisiting the Rolling Window, 33.
2. Fama, E. F., & French, K. R. (2002). The Equity Premium. The Journal of Finance, 57(2), 637–659. https://doi.org/10.1111/1540-6261.00437
3. Hsiao, C. Y.-L., & Morley, J. (2015). Debt and Financial Market Contagion. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2562121
4. Narayan, P. K., & Bannigidadmath, D. (2015). Are Indian stock returns predictable? Journal of Banking & Finance, 58, 506–531. https://doi.org/10.1016/j.jbankfin.2015.05.001
5. Polk, C., Thompson, S., & Vuolteenaho, T. (2006). Cross-sectional forecasts of the equity premium☆. Journal of Financial Economics, 81(1), 101–141. https://doi.org/10.1016/j.jfineco.2005.03.013
6. Rapach, D. E., Strauss, J. K., & Zhou, G. (n.d.). Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, 62.
7. Welch, I., & Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. Review of Financial Studies, 21(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014
8. White, L. J. (2010). Markets: The Credit Rating Agencies. Journal of Economic Perspectives, 24(2), 211–226. https://doi.org/10.1257/jep.24.2.211

Appendix

Table 1: Summary Statistics

The table provides the Tukey 5 number summary, along with the mean, third and fourth order sample moments of NIFTY 500 TRI returns, various proxies for the risk-free rate, 10-year Government bond yields and 10-year BBB Corporate bond yields where the rate in the table are on a monthly basis[9]. The period of data for all variables is from June 2000 till March 2018, except for 10-year BBB Corporate bond*.

	Min	1st Quartile	Median	Mean	3rd Quartile	Max	Skewness	Kurtosis
NIFTY 500 TRI Returns	-27.18%	-2.53%	1.50%	1.44%	65.74%	5.66%	-0.17	2.89
Call Money	0.06%	0.46%	0.55%	0.56%	0.66%	1.27%	0.42	2.03
30-day T-bill	0.24%	0.46%	0.54%	0.55%	0.66%	0.98%	-0.01	-0.21
3-month T-bill	0.27%	0.46%	0.56%	0.56%	0.67%	0.93%	-0.02	-0.45
1-year T-bill	0.30%	0.48%	0.58%	0.58%	0.67%	0.91%	0.02	0.50
10-year G-Bond	0.42%	0.58%	0.64%	0.64%	0.69%	0.99%	0.59	1.48
BBB Corp Bond	0.77%	0.93%	0.97%	0.97%	1.00%	1.16%	-0.13	0.69

[9] All figures reported in the table are monthly figures. E.g., mean Call Money rate is 0.56%. For an annual estimate, we multiply by 12, i.e., 0.56 X 12 = 6.72%.

* Period for 10-year BBB Corporate Bond is July 2004 to March 2018 due to constraints on data availability.

Table 2

Panel A: **Average returns of NIFTY 500 Total Returns Index**

This table gives the average NIFTY 500 returns estimates for the period of June 2000 – March 2018 computed using arithmetic average, geometric average and on a continuously compounded basis. The formulas for computation of these measures are shown alongside.

Arithmetic Average	17.25%
Geometric Average	15.07%
Continuous Compounding	14.04%

$$Arithmetic\ Average \text{ Returns} = \frac{\sum_{t=1}^{k}\left(\frac{(Index\ Level)_t}{(Index\ Level)_{t-1}} - 1\right)}{k} \times 12$$

$$\text{Geometric Average Returns} = \left(\prod_{t=1}^{k}\left(\frac{(\text{Index Level})_t}{(\text{Index Level})_{t-1}}\right)\right)^{\frac{12}{k}} - 1$$

$$\text{Continuous Compounded Average Returns} = \frac{\log\left(\frac{\text{Index Value}_k}{\text{Index Value}_1}\right)}{k} \times 12$$

where $(\text{Index Level})_t$ denotes the TRI index level at the end of month t and k is the number of months from June 2000 to March 2018.

Panel B: Realized Annual ERP from 2000-2018

Annual ERP Estimates with various proxies for risk-free rates and average ERP for the period June 2000 – March 2018. Risk-free rate corresponding to each risk-free rate proxy is the arithmetic average of daily observed, annualized rates of the risk-free rate proxy. ERP is calculated as the difference of TRI Returns and Risk-free Rate corresponding to each risk-free rate proxy. The ERP is calculated for all the three TRI return estimates by subtracting the corresponding risk-free rate proxy from the TRI returns.

Proxy	Risk-Free Rate (Arithmetic Average)	ERP (Arithmetic Average)	ERP (Geometric Average)
Call Money	6.66%	10.59%	7.67%
T-Bill 30d	6.56%	10.70%	7.78%
T-Bill 3m	6.71%	10.55%	7.62%
T-Bill 1y	6.91%	10.35%	7.41%

Table 3: **ADF test for ERP and Predictor Variables**

This table summarizes the Augmented Dickey-Fuller unit root test results for each of the dependent and predictor variables namely ERP, book to market, earnings to price, dividend yield, dividend price, dividend payout, stock variance, term spread and default spread.

Variable	**p-value**
Dependent Variable	
ERP (based on 30-day T-bill)*	0.01
Predictor Variable	
BOOK_MKT*	0.04
E/P*	0.03
DIV_PRICE	0.20
DIV_YIELD	0.43
DIV_PAYOUT	0.66
SVAR*	0.02
TERM_SPR	0.34
DEF_SPR	0.42

*No unit root at the 5% level of significance

Table 4: Number of Regressions for each predictor variable

The following table gives the number of regressions for each predictor variable using each of the regression methods. The number is same for both, rolling and augmented regression methods. The number of regressions vary across predictor variables because of different starting period from which its data is available.

Predictor Variable	Number of Regressions
BOOK_MKT	121
E/P	137
DIV_PRICE	136
DIV_YIELD	136
DIV_PAYOUT	136
SVAR	161
TERM_SPR	160
DEF_SPR	111

Table 5: Ranking in Relative Group: ΔRMSE and R^2 using Rolling Regression

The table presents statistics of out-of-sample prediction of ERP done through rolling regression method. Results of ΔRMSE and R^2 for each of the predictor variables and the relative ranking of the variables based on both the measures is shown in the table. ΔRMSE and R^2 are calculated with respect to the predictions based on historical mean ERP. We also include an equally weighted average rank measure to compare the overall predictive power. The numbers shown under ΔRMSE are in percent per month while the numbers under R^2 are simple percentages for the overall out-of-sample period.

Predictor	**ΔRMSE**	**Rank (ΔRMSE)**	**R^2 (%)**	**Rank (R^2)**	**Composite Rank**
BOOK_MKT	-0.0023	6	-6.83	7	6
E/P	-0.0027	8	-7.66	8	8
DIV_PRICE	-0.0024	7	-6.73	6	6
DIV_YIELD	-0.001	4	-2.91	4	4
DIV_PAYOUT	0.0004	1	1.18	1	1
SVAR	-0.0012	5	-3.35	5	5
TERM_SPR	-0.0059	10	-17.67	10	10
DEF_SPR	0.0034	9	-12.44	9	9
MEAN_COMB	-0.0008	3	-2.24	3	3
HIST_MEAN	0	2	0	2	2

Table 6: Ranking in Absolute Group: MAPE and PSCP using Rolling Regression

The table reports statistics of out-of-sample prediction of ERP using rolling regression method. Results of MAPE and PSCP for each of the predictor variables and the relative ranking of the variables based on both the measures is shown in the table. We also include an equally weighted average rank measure to compare the overall predictive power. The numbers shown under MAPE and PSCP are simple percentages for the overall sample period.

Predictor	MAPE (%)	Rank (MAPE)	PSCP (%)	Rank (PSCP)	Composite Rank
BOOK_MKT	165.8	5	51.24	8	6
E/P	176.71	8	51.82	7	8
DIV_PRICE	173.49	7	47.79	10	9
DIV_YIELD	160.66	3	52.94	6	5
DIV_PAYOUT	160.28	2	54.41	4	2
SVAR	171.52	6	57.76	1	4
TERM_SPR	178.89	9	53.75	5	7
DEF_SPR	192.86	10	50.45	9	10
MEAN_COMB	141.24	1	57.14	2	1
HIST_MEAN	164.66	4	57.14	2	2

Table 7: Sensitivity of results based on rolling regression method for varying in-sample period lengths

The table presents ΔRMSE of the predictor variables of ERP using rolling regression method based on in-sample period of 36 ,48 and 60 months.

Predictor	ΔRMSE 36-month	Rank	ΔRMSE 48-month	Rank	ΔRMSE 60-month	Rank
BOOK_MKT	-0.0027	8	-0.0023	6	-0.0038	9
E/P	-0.0038	9	-0.0027	8	-0.0032	7
DIV_PRICE	-0.0002	4	-0.0023	7	-0.0012	3
DIV_YIELD	-0.0019	7	-0.0010	4	-0.0009	2
DIV_PAYOUT	-0.0001	3	0.0004	1	-0.0033	8
SVAR	-0.0012	5	-0.0012	5	-0.0025	6
TERM_SPR	-0.0059	10	-0.0059	10	-0.0053	10
DEF_SPR	-0.0017	6	-0.0034	9	-0.0017	5
MEAN_COMB	0.0000	1	-0.0008	3	-0.0013	4
HIST_MEAN	0.0000	2	0.0000	2	0.0000	1

Table 8: Ranking in Relative Group: ΔRMSE and R^2 using Augmented Regression

The table presents statistics of out-of-sample prediction of ERP using augmented regression method. Results of ΔRMSE and R^2 for each of the predictor variables and the relative ranking of the variables based on both the measures is shown in the table. ΔRMSE and R^2 are calculated with respect to the predictions based on historical mean ERP. We also include an equally weighted average rank measure to compare the overall predictive power. The numbers shown under ΔRMSE are in percent per month while the numbers under R^2 are simple percentages for the overall out-of-sample period.

Predictor	ΔRMSE	Rank (ΔRMSE)	R^2 (%)	Rank (R^2)	Composite Rank
BOOK_MKT	-0.0021	6	-6.21	6	6
E/P	-0.0028	8	-8.02	8	8
DIV_PRICE	-0.002	4	-5.63	4	4
DIV_YIELD	-0.002	5	-5.68	5	5
DIV_PAYOUT	-0.0026	7	-7.47	7	7
SVAR	-0.003	10	-8.89	10	10
TERM_SPR	-0.003	9	-8.79	9	9
DEF_SPR	-0.0011	2	-3.96	2	2
MEAN_COMB	-0.0017	3	-5.08	3	3
HIST_MEAN	0	1	0	1	1

Table 9: Ranking in Absolute Group: MAPE and PSCP using Augmented Regression

The table presents statistics of out-of-sample prediction of ERP done through augmented egression method. Results of MAPE and PSCP for each of the predictor variables and the relative ranking of the variables based on both the measures is shown in the table. We also include an equally weighted average rank measure to compare the overall predictive power. The numbers shown under MAPE and PSCP are simple percentages for the overall out-of-sample period.

Predictor	MAPE (%)	Rank (MAPE)	PSCP (%)	Rank (PSCP)	Composite Rank
BOOK_MKT	193.44	6	54.55	5	5
E/P	341.50	10	47.45	10	10
DIV_PRICE	237.18	9	57.35	3	6
DIV_YIELD	223.46	8	54.41	6	8
DIV_PAYOUT	213.95	7	51.47	7	8
SVAR	148.79	2	55.28	4	2
TERM_SPR	125.07	1	47.5	9	4
DEF_SPR	189.67	5	47.75	8	7
MEAN_COMB	161.41	3	59.01	1	1
HIST_MEAN	167.57	4	57.76	2	2

Table 10: Variation in Ranks Across Groups

The table presents the relative ranking of predictor variables under rolling and augmented regression methods. Ranks under Relative Group (relative ranking measures) and Absolute Group (absolute ranking measures) are composite rankings for each of the variables under the Groups. In Relative Group, variables are ranked based on ΔRMSE and R^2 while in Absolute Group, variables are ranked based on MAPE and PSCP. The last two columns of the table report the mean rank and variation in ranks of the variables across both Groups and regression methods.

	Rolling Regression		Augmented Regression			
Predictor	**Relative Group Rank**	**Absolute Group Rank**	**Relative Group Rank**	**Absolute Group Rank**	**Mean Rank**	**Variance in Ranks**
BOOK_MKT	6	6	6	5	6	0.19
E/P	8	8	8	10	10	0.75
DIV_PRICE	6	9	4	6	7	3.19
DIV_YIELD	4	5	5	8	5	2.25
DIV_PAYOUT	1	2	7	8	3	9.25
SVAR	5	4	10	2	4	8.69
TERM_SPR	10	7	9	4	9	5.25
DEF_SPR	9	10	2	7	8	9.5
MEAN_COMB	3	1	3	1	2	1
HIST_MEAN	2	2	1	2	1	0.19

Table 11: Comparison of Augmented and Rolling regression methods

The table compares the difference in the various measures of prediction accuracy across both regression methods. For each predictor variable, we compute the difference between the value of accuracy measure observed under augmented regression and that observed under rolling regression. The numbers shown under MAPE and PSCP are simple percentages for the out-of-sample period.

Augmented - Rolling Regression		
Predictor Variable	**MAPE (%)**	**PSCP (%)**
BOOK_MKT	27.64	3.31
E/P	164.79	-4.37
DIV_PRICE	63.69	9.56
DIV_YIELD	62.8	1.47
DIV_PAYOUT	53.67	-2.94
SVAR	-22.73	-2.48
TERM_SPR	-53.82	-6.25
DEF_SPR	-3.19	-2.7
MEAN_COMB	20.17	1.87
HIST_MEAN	2.91	0.62
Average[+]	31.59	-0.19
Count[++]	7/10	5/10

[+] We observe from sign of the *Average* value of MAPE and PSCP, that rolling regression method, on average, gives more accurate ERP predictions than the augmented regression method.

[++] We also observe from *Count* that 1) 7 out of 10 predictors perform better in the rolling regression method than in the augmented regression method under the MAPE measure and 2) 5 out of 10 predictors perform better in the rolling regression method than in the augmented regression method under the PSCP measure.