[The Indonesian Capital Market Review](https://scholarhub.ui.ac.id/icmr)

Volume 12
Number 1 *January*

Article 4

1-30-2020

Determinants and Prediction Accuracy of Price Multiples for South East Asia: Conventional and Machine Learning Analysis

Himanshu Joshi
FORE School of Management (India), himanshu@fsm.ac.in

Rajneesh Chauha
FORE School of Management (India), rajneesh@fsm.ac.in

Indonesian Capital Market Review 12 (2020) 42-54

INDONESIAN
CAPITAL MARKET REVIEW

Determinants and Prediction Accuracy of Price Multiples for South East Asia: Conventional and Machine Learning Analysis

Himanshu Joshi* and Rajneesh Chauhan
FORE School of Management, India

(Received: January 2020 / Revised: February 2020/ Accepted: May 2020 / Available Online: June 2020)

The present study evaluates determinants of price multiples and their prediction accuracy using ordinary least square (OLS) regression and machine learning-based shrinkage methods for the South East Asian markets. Price multiples examined in the research are price to earnings (P/Es), price to book (P/B), and price to sales (P/S). Data has been collected from Thomson Reuters Eikon. The study recommends that the P/B ratio is the best price multiple for developing a price-based valuation model. Beside fundamental determinants of the multiple, various firm-level control variables, namely, firm size, cash holding, strategic holding, stock price volatility, firms' engagement in Environment, Social, and Governance (ESG) activities, dividend yield, and net profit margin impact firm's P/B multiple. Positive coefficients of consumer non-cyclical and healthcare dummies indicate a preference for defensive stocks by the investors. Application of machine learning-based shrinkage methods ensures the accuracy of prediction even with out-of-sample forecasting.

Keywords: *Price multiples, South East Asia, ridge regression, lasso, shrinkage method*

JEL Classification: C88, G12, G30, G35

Introduction

The relative valuation is based on the principle that comparable assets should be priced similarly by the market. The process of relative valuation involves three essential steps: first, finding the comparable assets priced by the market; second, scaling the market prices to a common variable to generate uniform comparable multiples, and third, adjusting for the differences across assets when comparing their uniform multiples (Damodaran, 1996). When applied to equity valuation, relative valuation primarily relies on multiples based on either market price of the stock or the value of the enterprise. To standardize the market prices of comparable stocks and values of comparable enterprises, they are divided by earnings available to equity shareholders and the aggregate firm-level earnings (EBDITA), respectively. In practice, comparable firms are generally selected from the line of business of the firm being valued. If there are an adequate number of firms available in the industry for comparison, the list is trimmed further using other scales, like the size of the firm or strategic holdings. Alternatively, firms can be grouped in terms of valuation fundamentals like firms' regression

* Corresponding author's email: himanshu@fsm.ac.in

beta, potential growth in earnings, and ROE. Regardless of how prudently a list of comparable firms is prepared, there will always be a possibility of differences in firm characteristics. Usually, subjective adjustments are made to control these differences. Alternatively, multiples can be modified to consider the most important variable determining the multiple. With modified multiples, firms become comparable in all other aspects of value, other than the one being controlled for. Subjective adjustments and modified multiples do not yield desired results when the relationship between multiples and expounding variables becomes complex. Sector regression can be used to overcome the limitations of subjective adjustments and modified multiples. In sector regression, price and value multiples are regressed with fundamental independent variables (earnings per share [EPS], growth rate, payout, and firm beta) to explain the differences across firms. The results of the sector regression provide a measure of the relationship between the multiple and the explanatory variables being used. Contrasting the modified multiple approaches, sector regression allows controlling more than one variable and also considers cross effects across these variables. The sector regression approach uses only the fundamental variables related to the multiple. The sector regression can be restrictive when an adequate number of comparable firms are not available in a particular sector/industry. In such cases, the market regression approach is appropriate. In the market regression approach, comparable firms are selected from the entire market cutting across various sectors and industries. The market regression approach not only focuses on the fundamental factors explaining the relative multiples but also attempts to improve the explanatory power of the regression model by adding more and more independent variables that can explain the differences among the comparable firms. In the present study, we attempt to evaluate determinants of price multiples and their prediction accuracy using a market regression approach for firms listed in the South East Asian market. With increasing complexity in market and business models of the firms, a large number of firm-level determinants like firm size, ownership structure, cash holding, etc. have acquired descriptive power to explain the price multiples. Therefore, in addition to the fundamental variables, we have used many firm-level control variables in market regression analysis. Machine learning-based regression methods like Ridge Regression and Lasso are also used to overcome the limitations of conventional regression analysis.

Literature Review

Several textbooks on corporate finance, financial economics, and valuation extensively discuss the valuation multiples based on stock price as well as firms' value (Copeland, Koller & Murrin, 1994; Damodaran, 1996; Kasper, 1997; Healey, Palepu & Bernard, 2000; Pinto, et al., 2016). Besides the textbooks, several research papers also discuss the efficacy of various price and value multiples in firm valuation and provide an account of their prediction accuracy.

Boatsman & Baskin (1981) provided empirical pieces of evidence regarding the predictive accuracy of price to earnings (P/E) ratios. They used two different sets of firms belonging to the same industry for their analysis. The authors observed that valuation errors can be minimized by selecting comparable firms based on their analogous historical earnings growth rates. LeClair (1990) tested the P/E valuation method by selecting comparable firms based on industry classification. He used three measures of earnings: current earnings, last years' average earnings, and earnings on tangible and intangible assets. Findings of LeClair suggest that average earnings perform best for the valuation model. Alford (1992) studied the accuracy of the P/E valuation method by selecting comparable firms based on three-fold criteria, namely, industry classification, risk, and earnings growth rate. The accuracy of the P/E valuation method for each method of selected comparable firms was estimated by comparing each firm's predicted stock price with its observed price. Alford found that industry classification or a combination of risk and earnings growth rates are effective criteria for selecting comparable firms. Results also suggested that segregating industries by risk or growth rate does not improve accuracy. Also, adjusting the P/E ratio for variation in leverage across comparable firms results in reduced accuracy.

Penman (1997) combined P/E and price to book (P/B) multiples to use the information provided by both the multiples in the prediction of the stock price. He experimented by assigning different weights to two multiples and found that weights vary in a nonlinear way over the amount of earnings relative to book value and systematically over time. Estimated weights were found to be robust over time and appropriate for out of sample forecasting. Tasker (1998) reported the systematic use of industry-specific multiples, suggesting the suitability of different multiples for different sectors and industries. Baker & Ruback (1999) compared the relative performance of industry multiples based on Earnings before interest, tax, depreciation, and amortization (EBITDA), earnings before interest and taxes (EBIT), and sales. They reported that absolute valuation errors were proportionate to the value and harmonic mean of industry multiples were close to the Monte Carlo simulations based minimum-variance.

Liu, Nissim, & Thomas (2002a) examined the valuation performance of a wide-ranging list of value drivers to determine their appropriateness for stock price prediction. They also examined variation in performance of multiples across industries and over time by using alternative definitions of multiples (using forward or current earnings to estimate the multiple). Authors reported that forward earnings perform the best and the intrinsic value measures based on residual earnings models perform worse. The performance of forward earnings-based multiples improves with the length of prediction time. They also reported that among the drivers using historical data, earnings perform the best, followed by book value, cash flows, and sales, respectively. Liu, Nissim, & Thomas (2002b) extended their work and examined the stock price prediction performance of industry multiples using data from ten countries. Their findings were analogous to their single country results. They reported that multiples based on earnings perform the best, followed by dividends, and cash flows. The worst performance was recorded for sales-based multiples. Liu, Nissim, & Thomas (2007) compared the prediction accuracy of cash flows based price multiples with earnings based price multiples. They found that regardless of intuitive understanding that operating cash flows provide better summary measures of value, reported earnings to explain stock price better than the estimated cash flows.

Huang, Tsai, & Chen (2007) examined the P/E multiples by decomposing it into fundamental and residual components. They found that P/E multiples are explained by firm-specific as well as macroeconomic factors. Forecasted long-term growth rate, dividend payout ratio, and firm size were found to have a positive association with P/E multiples, while risk and aggregate bond yields were having a negative association with the multiples.

Sehgal & Pandey (2010) evaluated alternative price multiples for equity valuation purposes using data from 145 large Indian firms. The authors generated price forecasts for four price-based multiples by regressing the observed historical prices on various value drivers. Price multiples used in the study were P/Es, P/B price to cash flow, and P/S. Forecast accuracy of different multiples was measured using root mean squared error (RMSE) and Theil's coefficient. They found that P/E multiples provide the best price forecast compared to the other three price multiples. They also experimented with price forecasts based on pairwise combinations of these price multiples. The value driver combination of book value-sales was found to be the most efficient in terms of error minimization. Nevertheless, P/E as a standalone multiple performed better in equity valuation as compared to all the combinations of value drivers. Pereira, Basto, & Ferreira da Silva (2016) examined corporate failure prediction using logistic Lasso and Ridge Regression. The results showed that the Lasso and Ridge models tend to favor the category of the dependent variable that appears with heavier weight in the training set when compared to the stepwise methods.

Table 1. Fundamental determinants for price multiples

Multiple	Fundamental Determinants
P/E multiple	Projected growth rate, payout, risk
P/B multiple	Projected growth rate, payout, risk, return on equity
P/S multiple	Projected growth rate, payout, risk, net profit margin

Table 2. Definitions of value determinants and other firm-level control variables

Variable	Definition
Projected Growth Rate	Thomson Reuters
Payout	Dividend per share/EPS
Risk	Five-year monthly beta
ROE	Net income/shareholders' equity
Net Profit Margin	Net income/total revenues
Firm Size	Log of total assets
Dividend Yield	Dividend per share/stock price
Volatility	Annualized standard deviation of the relative price change for the most recent 200 trading days
Cash Holding	Cash plus marketable securities/total assets
Strategic holding	% of strategic ownership
Return on Capital Employed	EBIT/(total assets – current liabilities)
ESG Dummy	Takes the value of 1 if the firm has ESG score, otherwise, it takes the value of 0.
Sector Dummy	Takes the value of 1 if the firm belongs to a particular sector, otherwise, it takes the value of 0.

No particular study is available on the accuracy and determinants of price multiples in the context of South East Asian countries. The present research evaluates the prediction accuracy of price multiples across various sectors for South East Asian firms and attempts to identify the fundamental drivers for these multiples. This paper extends the work of Huang, Tsai, & Chen (2007) by considering three price multiples, namely, P/E, P/B, and P/S. We have classified determinants of these multiples into two groups: fundamental value determinants and other firm-level control variables. Fundamental value determinants are expected growth rate, payout, risk, ROE, and net profit margin. Firm-level control variables used in the study are firm size, dividend yield, stock price volatility, cash holding, strategic holding, return on capital employed (ROCE), and a dummy for the environment, social, and governance (ESG). The industry classification of the Thomson Reuters Eikon database has been used. The present work contributes to the South East Asian market literature by evaluating the key price multiples and their fundamental value drivers. It also identifies various firm-level determinants of price multiples using conventional multiple regression analysis as well as the shrinkage regression method, namely, Ridge Regression and Lasso.

Research Methodology

Sample

Data for 842 firms from South East Asian countries, namely, Malaysia, Indonesia, the Philippines, Thailand, Vietnam, and Singapore have been collected from Thomson Reuters' Eikon database. There are 74 firms from basic material, 129 from consumer cyclical, 106 from consumer non-cyclical, 37 from energy, 221 from financials, 30 from healthcare, 151 from industrial, 40 from technology, 19 telecommunications, and 35 from utility sectors. Three price multiples, namely, P/E, P/B, and P/S are used as the dependent variable in the study. Fundamental determinants for price multiples used in the study is shown in table 1.

Several control variables having explanatory power for the respective price multiples have been used in regression in addition to the key-value determinants. These variables are firm size, dividend yield, 200-days stock price volatility, cash holding, strategic holding, net profit margin, return on equity (ROE), and ROCE (Table 2). A dummy variable for firms' engagement in ESG practices has been used. To ascertain the fact of whether sector variation impacts firms' price multiples, N-1 sector dummies are used. These dummy variables are used for the sectors taken in the study except financial sector firms.

Model Specification

Linear multiple regression analyses have been conducted using price multiples as the dependent variables and their respective fundamental determinants as independent variables. Several firm-level control variables are used as control variables. Dummy variables are used for ESG and sectoral classification. The dummy variable of ESG takes a value of 1 if the firm has an ESG score in Thomson Reuters' Eikon database; otherwise, it takes the value of a zero. Similarly, sectoral dummy takes the value of 1 if the firm belongs to a particular sector, otherwise, it takes the value of zero. Fundamental determinants and other control variables are defined in Table 2.

$$\begin{aligned} P/Es = {} & \beta_0 + \beta_1 \text{ (Growth Rate)} \\ & + \beta_2 \text{ (Payout Ratio)} + \beta_3 \text{ (Risk)} \end{aligned} \quad (1)$$

$$\begin{aligned} P/Es = {} & \beta_0 + \beta_1 \text{ (Growth Rate)} \\ & + \beta_2 \text{ (Payout Ratio)} + \beta_3 \text{ (Risk)} \\ & + \beta_4 \text{ (Firm Size)} + \beta_5 \text{ (Div\_Yield)} \\ & + \beta_6 \text{ (Volatility)} + \beta_7 \text{ (Cash Holding)} \\ & + \beta_8 \text{ (Strategic Holding)} \\ & + \beta_9 \text{ (Net Profit Margin)} + \beta_{10} \text{ (ROE)} \\ & + \beta_{11} \text{ (ROCE)} + \beta_{12} \text{ (ESG\_Dummy)} \\ & + \beta_{13} \text{ (Sector\_Dummy)}_{N-1} \end{aligned} \quad (2)$$

$$\begin{aligned} PB = {} & \beta_0 + \beta_1 \text{ (Growth Rate)} \\ & + \beta_2 \text{ (Payout Ratio)} + \beta_3 \text{ (Risk)} \\ & + \beta_4 \text{ (ROE)} \end{aligned} \quad (3)$$

$$\begin{aligned} PB = {} & \beta_0 + \beta_1 \text{ (Growth Rate)} \\ & + \beta_2 \text{ (Payout Ratio)} + \beta_3 \text{ (Risk)} \\ & + \beta_4 \text{ (ROCE)} + \beta_5 \text{ (Firm Size)} \\ & + \beta_6 \text{ (Div\_Yield)} + \beta_7 \text{ (Volatility)} \\ & + \beta_8 \text{ (Cash Holding)} \\ & + \beta_9 \text{ (Strategic Holding)} \\ & + \beta_{10} \text{ (Net Profit Margin)} \\ & + \beta_{11} \text{ (ROE)} + \beta_{12} \text{ (ESG\_Dummy)} \\ & + \beta_{13} \text{ (Sector\_Dummy)}_{N-1} \end{aligned} \quad (4)$$

$$\begin{aligned} P/S = {} & \beta_0 + \beta_1 \text{ (Growth Rate)} \\ & + \beta_2 \text{ (Payout Ratio)} + \beta_3 \text{ (Risk)} \\ & + \beta_4 \text{ (NP Margin)} \end{aligned} \quad (5)$$

$$\begin{aligned} P/S = {} & \beta_0 + \beta_1 \text{ (Growth Rate)} \\ & + \beta_2 \text{ (Payout Ratio)} + \beta_3 \text{ (Risk)} \\ & + \beta_4 \text{ (Firm Size)} + \beta_5 \text{ (Div\_Yield)} \\ & + \beta_6 \text{ (Volatility)} + \beta_7 \text{ (Cash Holding)} \\ & + \beta_8 \text{ (Strategic Holding)} \\ & + \beta_9 \text{ (Net Profit Margin)} \\ & + \beta_{10} \text{ (ROE)} + \beta_{11} \text{ (ROCE)} \\ & + \beta_{12} \text{ (ESG\_Dummy)} \\ & + \beta_{13} \text{ (Sector\_Dummy)}_{N-1} \end{aligned} \quad (6)$$

To improve the prediction accuracy and regression models' fit, coefficient estimates of the regression outcome can be shrinked towards zero. These shrinkage techniques improve the model fit by significantly reducing their variance. Two regression coefficients shrinking techniques used in the present study are Ridge Regression and Lasso. In Ridge Regression coefficients are estimated by minimizing the following:

$$\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}\beta_j^2 = RSS + \lambda\sum_{j=1}^{p}\beta_j^2 \quad (7)$$

Where $\lambda \geq 0$ is a tuning parameter to be determined separately.

Ridge Regression technique is a shrinkage method that involves fitting a model involving all p predictors. However, unlike the conventional least-square method, in the Ridge Regression technique estimated coefficients are shrunken towards zero relative to the least square estimates. This shrinkage has the effect of reducing variance. The shrinkage method can also be performed for variable selection. However, Ridge Regression always generates a model involving all the predictors. The penalty $\lambda \sum_{j=1}^{p}\beta_j^2$ in equation (7) will shrink all of the coefficients towards zero, but it will not set any of them exactly to zero. Therefore, Ridge Regression is appropriate for out-of-sample prediction accuracy, but it can create a challenge in model interpretation in scenarios in which several predictors (dependent variables) are relatively large. In the present research, besides the fundamental determinants of the multiples, there are a large number of firm-level control variables that can impact the value of respective multiples. Therefore, a model that can reduce the number of predictors by forcing some of the coefficient estimates to be exactly zero will improve the interpretability of the model.

The Lasso is an alternative technique to Ridge Regression, which overcome this limitation. In Lasso, the penalty has the effect of forcing some of the coefficient estimates to be exactly zero when the tuning parameter is sufficiently large. Therefore, Lasso performs the task of variable selection. As a result, models generated from the Lasso are generally much easier to interpret than those produced by Ridge Regression. The Lasso coefficients, β^{L}_{λ}, minimizes the quantity:

$$\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2+\lambda\sum_{j=1}^{p}|\beta_j|=RSS+\lambda\sum_{j=1}^{p}|\beta_j| \quad (8)$$

Comparing equations for Ridge Regression (7) and Lasso (8) reveals that the β_j^2 in the Ridge Regression penalty has been replaced by $|\beta_j|$ in the Lasso penalty. Otherwise, both have similar formulations. The empirical findings are based on RMSE and adjusted R-squared. Prediction accuracy of multiples for ordinary least square (OLS) regression, Ridge Regression, and Lasso has been assessed using two measures, namely, adjusted R^2 and RMSE. Adjusted R-squared does not automatically increase when additional variables are added to the regression; it is adjusted for degrees of freedom. In fact, in the addition of a new variable, adjusted R-squared can decrease if adding that variable results in only a small increase in R-squared.

$$Adjusted\ R^2 = 1-\frac{n(n-1)}{(n-k-1)} * (1-R^2) \quad (9)$$

Root mean squared error is the standard deviation of the residuals of the regression model. Residuals measure the distance of the regression line from the observed data points. Therefore, RMSE is a measure of how concentrated the data is around the line of best fit.

$$RMSE = \sqrt{\Sigma_{i=1}^{n}\frac{(\hat{y}-y)}{n}} \quad (10)$$

Results and Discussions

Table 3 presents the results of the multiple regression analysis for three price multiples, namely, P/E, P/B, and P/S. These results are heteroscedasticity consistent. Columns (1), (3), and (5) presents the result of regression using only fundamental determinants as independent variables, while columns (2), (4), and (6) present the results considering all the firm-level control variables in regression analysis.

Price to earnings multiple regression with only key drivers as independent variables (column 1) has adjusted R-squared of 0.0348. Out of three key drivers, only the EPS growth rate is statistically significant and has a negative coefficient. A negative coefficient indicates that the firm having higher EPS growth rates have a lower value of P/E multiple. For P/E multiples using all the firm-level control variable (column 2), adjusted R-squared improves to 0.1424. Besides the growth rate, dividend yield, firm's cash holdings, net profit margin, and ROE have statistically significant coefficients. Out of these statistically significant variables, only the dividend yield has a negative coefficient, while all other variables have positive coefficients. This means that the firm's cash holdings, net profit margin, and ROE influence a firm's P/Es multiple positively. Out of the sectoral dummies, basic materials, consumer cyclicals, consumer non-cyclicals, and healthcare have statistically significant positive coefficients, which indicates that firms from these sectors have a higher value of P/Es multiples in comparison to the absent sectoral dummy, i.e. finance sector.

Adjusted R-squared for another price multiple, P/B is 0.6548, which means that key drivers explain only 65.48% of P/B (column 3). Out of the key drivers, payout ratio, growth rate, and ROE have statistically significant coefficients. The positive coefficient of payout ratio indicates that South East Asian firms paying higher dividends as a percentage of their earnings command higher P/B ratio. Similarly, ROE has a positive coefficient indicating that firms capable of generating a surplus for equity holders command a higher P/B ratio. Contrary to the theoretical assumption, the growth rate

Table 3. Heteroscedasticity-consistent regression results for P/Es, P/B, and P/S for South East Asian Firms

	P/E	P/E	P/B	P/B	P/S	P/S
Adjusted R Square	0.0348	0.1424	0.6548	0.6954	0.3903	0.4313
RMSE	15.4261	14.3192	2.3878	2.2146	3.2088	3.0671
F – Statistics	8.2001	5.7304	24.9685	92.4646	13.6647	31.3800
C	23.3308	8.8413	1.4869	5.7302	0.7139	6.8646
	6.5286	0.4660	8.0557	2.6275	1.1984	3.0572
Growth Rate	−57.9323	−64.7181	−22.9937	−21.1743	0.1722	−0.7324
	−3.6801***	−3.7195***	−8.0917***	−7.7352***	0.1984	−0.5468
Pay Out	−4.5889	0.2572	−1.2166	−0.7940	1.3570	1.5115
	−1.4826	0.0794	−4.7455***	−3.3049***	2.2131**	2.3654**
Risk	−0.5022	−0.1765	0.0355	0.2173	0.2310	0.3579
	−0.2861	−0.0918	0.2324	1.5690	0.4775	0.7471
ROE		9.5957	23.2564	21.0541		1.2577
		1.8670*	8.9846***	8.1298***		1.1051
Net Profit Margin		3.4993		−0.06547	7.2025	7.2060
		2.2578**		−0.2921	2.3114**	2.1299**
Firm Size		0.3542		−0.1925		−0.2179
		0.5102		−2.3596**		−2.004**
Dividend Yield		−137.3333		−13.3819		−16.0768
		−4.7123***		−1.9817**		−2.0147**
ESG Dummy		−0.3557		0.5695		0.5956
		−0.1707		1.6480*		1.5321
Volatility		0.0764		−0.0113		−0.0043
		0.9067		−1.8522*		−0.4259
Cash Holdings		8.8926		−1.2321		3.3157
		1.8299*		−1.5275		2.3435**
Strategic Holding		0.9151		0.9091		0.3865
		0.4355		2.5899***		0.7944
ROCE		−14.9921		7.6254		−2.8079
		−1.1573		2.1943**		−0.5940
Basic Material		11.1601		−0.0555		−1.4236
		1.7979*		−0.2078		−1.8294*
Consumer Cyclical		4.2589		0.1214		−1.5558
		2.3055**		0.5042		−1.9255*
Consumer Non-cyclical		7.3000		1.0355		−1.1902
		3.2101***		3.2212***		−1.4342
Energy		−1.0075		−0.1011		−1.9118
		−0.6536		−0.2865		−2.4096**
Healthcare		14.3874		1.5651		0.1227
		2.4142**		3.2668***		0.8834
Industrial		0.9618		−0.1139		−1.6545
		0.6946		−0.5637		−2.2508**
Technology		−0.7928		−0.4029		−1.7333
		−0.3786		−1.3819		−2.0577**
Telecommunication		3.2239		0.1610		1.0616
		1.4261		0.8198		0.8241
Utilities		2.2773		0.1681		−0.7298
		1.4686		0.5835		−1.0747

Source: Authors' Regression Result (2020)
Notes: *p < 0.10; **p < 0.05; ***p < 0.01

has a negative coefficient. The adjusted R-squared for the P/B ratio improves to 0.6954 on the inclusion of various firm-level control variables. Besides fundamental determinants, various firm-level control variables, namely, firm size, dividend yield, ESG dummy, stock price volatility, strategic holding, ROE, and ROCE have statistically significant coefficients. The growth rate, payout, firm size, dividend yield, and volatility have negative coefficients, while ESG dummy, strategic holding, ROE, and ROCE have positive coefficients. The negative coefficients of the growth rate and payout ratio collectively are contrary to the existing literature in finance. The negative coefficient of firm size indicates that smaller firms command high-

Table 4. Ridge Regression results for P/Es, P/B, and P/S for South East Asian Firms

	P/Es	P/B	P/S
RMSE	12.909	2.332	3.198
Adjusted R-Squared	0.157	0.562	0.499
C	20.473	4.843	2.741
Growth Rate	−34.284	−15.672	−0.237
Pay Out	0.583	−0.250	1.081
Risk	−0.200	0.104	0.148
Firm Size	−0.101	−0.174	−0.041
Dividend Yield	−84.838	−12.600	−8.176
ESG Dummy	0.397	0.418	0.056
Volatility	0.034	-0.010	−0.004
Cash Holdings	3.870	−1.334	1.613
Strategic Holding	1.869	1.069	−0.099
Net Profit Margin	0.325	−0.338	5.070
ROE	1.553	16.773	0.394
ROCE	−3.476	12.362	2.278
Basic Material	5.779	−0.158	−0.709
Consumer Cyclical	0.692	0.064	−0.828
Consumer Non-cyclical	3.574	1.054	−0.588
Energy	−2.546	−0.180	−1.216
Healthcare	7.735	1.443	0.361
Industrial	−0.964	−0.105	−0.861
Technology	−1.874	−0.357	−0.739
Telecommunication	1.195	0.725	1.305
Utilities	0.058	0.061	−0.245

Source: Authors' Regression Results (2020)

er P/B. The negative coefficient of dividend yield indicates that firms having higher dividend yields are considered as a surrogate to the bond market instruments, therefore they have lower P/B ratios. The negative coefficient of stock price volatility is in line with the existing literature, showing that the more volatile stocks are considered risky by the market participants and thus tend to have a lower P/B ratio. Out of the sectoral dummies, consumer non-cyclicals and healthcare have statistically significant and positive coefficients, indicating better P/B ratios for firms from these sectors in comparison to the financial sector firms.

Price to sales ratio (column 5) has adjusted R-squared of 0.3903, with only four fundamental determinants. Out of these fundamental determinants, only the payout ratio has a statistically significant value. This indicates that firms having higher dividend payout ratios have relatively higher P/S ratios. Adjusted R-squared for the P/S multiple improves considerably to 0.4313 on the inclusion of a firm-level control variable (column 6). Out of the firm-level control variables, firms' cash holdings and net profit margin have statistically significant positive coefficients, and firm size and dividend yields have statistically significant negative coefficients. Therefore, results indicate that firms with higher cash holdings and net profit margins command higher P/S multiples, while larger firms having higher dividend yields tend to have lower P/S multiples. According to the conventional multiple regression analysis, P/B seems to be the most appropriate price multiple and this multiple is better explained with the help of firm-level control variables such as firm size, dividend yield, ESG, volatility, strategic holding, and ROCE. Also, for P/B multiple significant sectoral variations have been recorded.

Table 4 gives the results of Ridge Regression for the three multiples, namely, P/Es, P/B, and P/S. Ridge Regression is a machine learning technique that shrinks the coefficients thus reducing the variance. The advantage of this method is not so much on improved interpretability but on improved prediction accuracy. In this technique, all the features are retained but the feature coefficients are reduced. Accordingly, in Table 4, all the coefficients are less than the coefficients that were deduced while using conventional OLS methods.

Table 5. Lasso Regression results for P/Es, P/B, and P/S for South East Asian Firms

	P/Es	P/B	P/S
RMSE	12.922	2.300	3.255
Adjusted R-Squared	0.156	0.562	0.555
C	21.607	5.566	1.246
Growth Rate	−38.844	−20.584	
Pay Out		−0.645	1.115
Risk		0.125	
Firm Size		−0.194	
Dividend Yield	−100.821	−13.006	−9.391
ESG Dummy		0.257	
Volatility		−0.008	
Cash Holdings		−0.860	1.071
Strategic Holding		0.879	
Net Profit Margin		−0.039	6.956
ROE		20.729	
ROCE		7.631	
Basic Material	5.586		
Consumer Cyclical		0.107	−0.141
Consumer Non-cyclical	2.936	1.047	
Energy		−0.033	−0.355
Healthcare	7.348	1.576	0.297
Industrial		−0.028	−0.226
Technology		−0.300	
Telecommunication		0.059	1.329
Utilities		0.060	

Source: Authors' Regression Result (2020)

In the case of regression of P/E multiples, amongst the fundamental value determinants growth rate had a substantial coefficient value albeit with a negative sign. This is in line with the results of conventional OLS regression where amongst all the fundamental value determinants only the growth rate was significant and again with a negative sign. Further, from amongst the firm-level control variables, dividend yield had the most substantial absolute coefficient values followed by cash holdings, ROCE, strategic holding, and ROE. Out of these, dividend yield and ROCE had a negative coefficient. Amongst the sectoral dummies, independent of the coefficient signs, healthcare, basic material, consumer non-cyclical, energy, technology, and telecommunication had the most substantial coefficients in the descending order. In the case of regression of P/B multiples, amongst the fundamental value determinants, growth rate and ROE had significant absolute coefficient values. This is in line with the results of conventional OLS regression where amongst all the fundamental value determinants growth rate, ROE had the most substantial effect on the value of P/B multiple. Further, from amongst the firm-level control variables, dividend yield and ROCE had the most substantial absolute coefficient values followed by cash holding and strategic holding. Out of these, dividend yield and cash holding had a negative coefficient. Amongst the sectoral dummies, healthcare and consumer non-cyclical had the most substantial coefficients in the descending order. In the case of regression of P/S multiple, amongst the fundamental value determinants, net profit margin and payout had substantial absolute coefficient values. This is in line with the results of conventional OLS regression where amongst all fundamental value determinants, the significant variables were net profit margin and payout. Thereafter, from amongst the firm-level control variables, dividend yield had the most substantial absolute coefficient values followed by ROCE and cash holding. Out of these, dividend yield had a negative coefficient. Amongst the sectoral dummies, independent of the coefficient signs, telecommunication, and energy had the most substantial coefficients in the descending order.

Table 5 presents the results of another shrinkage method, namely, Lasso for P/E, P/B, and

P/S multiples. In Lasso, the penalty has the effect of forcing some of the coefficient estimates to be exactly zero when the tuning parameter is sufficiently large. Therefore, Lasso performs the task of variable selection. As a result, models generated from the Lasso are generally much easier to interpret than those produced by Ridge Regression.

Both Ridge Regression and Lasso overcome the limitation of overfitting by shrinking the coefficients towards zero. However, Lasso forces some of the coefficients to be exactly equal to zero so that only highly significant independent variables are retained. For P/Es multiple, adjusted R-squared has improved marginally in comparison to the OLS regression, but it has remained almost the same as of Ridge Regression. For P/Es multiples, only one fundamental determinant growth rate has been retained by the Lasso method. The negative coefficient of the growth rate indicates that firms having a higher projected growth rate have a lower value of the multiple. When we compare the coefficient of the growth rate produced by OLS, Ridge Regression, and Lasso methods, it can be noticed that the Ridge Regression coefficient has less negative value than the OLS regression, but Lasso produces a more negative coefficient than the Ridge Regression. This confirms that the growth rate is one of the most significant determinants of the P/E multiple, with a strong negative influence on it. Firm-level determinant retained by the Lasso method is the dividend yield, again having a very strong negative coefficient, which indicates that firms having a higher dividend yield do not produce higher P/E multiples. Hypothetically, standalone negative coefficient of dividend yield is justified on the ground that stocks that generate higher dividend yields, are considered as surrogate to the bonds, and investor prefers to invest in the higher growth rate stocks rather than investing in higher dividend yield bonds. However, this relationship exists only when the macroeconomic environment is encouraging for business. In the case of economic slowdown or recession, investors prefer stocks with a better dividend yield. Nonetheless, the negative coefficient of dividend yield as well as the firm's growth rate is confounding. Coefficients of three sectoral dummies, namely, basic materials, consumer non-cyclicals, and healthcare have been retained in the Lasso method. As all these sectoral dummies have positive coefficients, it indicates that stocks belonging to these sectors enjoy better valuation in terms of P/E multiples. Moreover, all the three sectors for which coefficients have been retained by the Lasso are generally considered as defensive stocks by the investors. This indicates a preference for defensive stocks by the market participants. Since cross-section data have been employed for covering firms from various countries from South East Asia as well as different sectors, macro-economic variations across markets have not been the focus of the study. The South East Asian market has been considered as a cluster, having a contagious influence on the group members.

For P/B multiple, adjusted R-squared for the Lasso method is 0.562, which is the same as of the Ridge Regression, but lower than the value of OLS regression (0.695). This lower value of adjusted R-squared corresponds to the penalty imposed by Ridge and Lasso methods for an overfitting problem. All the fundamental determinants, as well as firm-level control variables, have been retained by the Lasso method for P/B multiple, indicating that every determinant proposed in the research has substantial explanatory power, be it firm size, cash holding or strategic holdings. Negative coefficients of the growth rate and payout again put forth the confounding result that the market prefers neither the high growth firms nor the high payout firms. Rather, a positive coefficient of risk, which is stock beta with their respective market indices, shows that a higher P/B multiple denotes a higher risk. Smaller firms and firms offering lower dividend yields tend to have a higher value of the multiple. Similarly, firms having lower cash holdings command a higher P/B. On the contrary, firms having engagements in environmental, social, and governance practices command a higher value of the multiple. Similarly, firms with a high concentration of strategic ownership enjoy a higher value of the multiple. Out of sectoral dummies, only the

https://scholarhub.ui.ac.id/icmr/vol12/iss1/4
DOI: 10.21002/icmr.v12i1.12051

Table 6. Prediction accuracy of multiples using OLS regression, Ridge regression, and Lasso

	OLS Regression		Ridge Regression		Lasso	
	Adjusted R^2	RMSE	Adjusted R^2	RMSE	Adjusted R^2	RMSE
P/Es	0.1424	14.3192	0.1570	12.909	0.157	12.922
P/B	0.6954	2.3146	0.5620	2.322	0.562	2.300
P/S	0.4313	3.0671	0.4990	3.198	0.555	3.255

Source: Authors' Regression Result (2020)

dummy for basic material has been removed, all other dummies have been retained, indicating that there is a substantial sectoral variation in P/B multiple's values. Sectoral dummies of consumer cyclicals, consumer non-cyclicals, healthcare, telecommunications, and utilities have positive coefficients, out of which healthcare and consumer non-cyclical have a relatively higher value of coefficients, which again confirms the investor's preference for defensive stocks. On the contrary, negative coefficients of energy, industrial, and technology stocks indicate relatively lower valuation for these sectors.

For P/S multiple, adjusted R-squared of Lasso regression is 0.555 which is a significant improvement over the R-squared value of 0.499 of Ridge Regression and 0.4313 of OLS regression. A higher value of adjusted R-squared for Ridge and Lasso over conventional OLS method indicates better goodness of fit along with the elimination of the in-sample model over-fitting problem. Payout is the only fundamental determinant retained by the Lasso method. A positive coefficient of payout indicates that firms paying a higher dividend as a percentage of their earnings, tend to command a higher value of P/S multiple. Other firm-level determinants retained by the Lasso are dividend yield, cash holdings, and net profit margin. The negative coefficient of dividend yield is analogous with the results of the other two multiples, which indicates that stocks offering higher dividend yields are considered surrogate to the bond market instruments and investors prefer high growth stocks for investments in comparison to high dividend yield stocks. The positive coefficient of cash holdings indicates that investors perceive higher cash holding as an indicator of enhanced product market control, which can influence the revenue generation capacity of the firm. The positive coefficient of net profit margin indicates that more profitable firms generate higher market capitalization on their revenue. Sectoral dummy variables retained by Lasso are healthcare, telecommunications, consumer cyclicals, energy, and industrials. Consistent with the results of the other two multiples, healthcare and telecommunication dummies have positive coefficients indicating favorable valuation by the market and consumer cyclicals, energy, and industrials having negative coefficients showing adverse valuation.

Table 6 presents the prediction accuracy data for three methods used in the study, namely, conventional OLS regression, Ridge Regression, and Lasso for three price multiples.

The coefficient of determination, adjusted R-square measures the fraction of the total variation in the dependent variable that is explained by the independent variables, while, RMSE is a measure of how concentrated the data is around the line of best fit. Therefore, higher the adjusted R-square, better the explanation of the dependent variable by independent variables, and lower the RMSE, better the in-sample forecasting accuracy of the regression model. Adjusted R-square has improved for P/E multiples from 0.1424 in the OLS method to 0.1570 in Ridge and Lasso methods, while RMSE diminishes from 14.3192 in OLS to 12.909 and 12.922 in Ridge Regression and Lasso, respectively. Ridge Regression and Lasso shrink the coefficient towards zero to overcome in-sample overfitting issues with the traditional OLS method. This is evident in the case of P/B multiple, where adjusted R-square got abridged from 0.6954 to 0.562 in Ridge and Lasso methods. Even the value of RMSE does not diminish for this multiple between OLS regression and shrinkage methods. However, the results of Ridge and Lasso are more reliable as they deliver better out of sample forecasting accuracy. For P/S multiple, adjusted R-square has improved from 0.4313 in OLS to 0.4990 and 0.555 in Ridge and Lasso methods, respectively, without any substantial change in the value of RMSE.

Conclusions

Overall P/B ratio seems to be the most appropriate price multiple to measure the valuation of firms using fundamental determinants and other firm-level control variables. It has the highest value of adjusted R-square and the smallest RMSE while working with only fundamental determinants. On inclusion of other firm-level control variables like firm size, ESG, cash holding, etc. adjusted R-square improves for the multiple in conventional OLS regression. Shrinkage methods—Ridge Regression and Lasso put penalty for in-sample overfitting and condense the adjusted R-square for delivering better out of sample forecasting. Lasso retains all the fundamental determinants as well as firm-level control variables introduced in the present research, demonstrating that every determinant proposed in the study has substantial explanatory power. Therefore, it can be concluded that beside the growth rate, payout, risk, and ROE which are fundamental determinants of P/B multiples, various firm-level control variables, namely, firm size, cash holding, strategic holding, stock price volatility, firms' engagement in environmental, social and governance initiative, dividend yield, and net profit margin impact firm's P/B multiple. Also, there is substantial sectoral variation in P/B multiple as all the sectoral dummies are retained by the Lasso method except basic material. P/S is the second-best multiple after P/B for the firm valuation. Besides payout, which is a fundamental determinant, firm-level determinants that explain P/S multiple are dividend yield and cash holding. P/Es multiple has been explained by the growth rate and dividend yield. Sectoral dummies of consumer non-cyclicals and healthcare have positive coefficients across all the three multiples, indicating a preference for defensive stocks by the investors.

Implications

The study recommends that investors and analysts shall use P/B multiple for relative valuation in the context in the South East Asian market. Application of machine learning-based shrinkage methods ensures the accuracy of prediction even with out-of-sample forecasting. For, firms having a negative value of P/B multiple, P/S multiple can be used. Positive coefficients of consumer non-cyclicals and healthcare dummies indicates a preference for defensive stocks by the investors.

The present work contributes to South Asian market investment and valuation literature by identifying the key price multiple, its determinants, and prediction accuracy using the conventional OLS regression method as well as machine learning-based shrinkage methods. P/B ratio emerges as the most appropriate valuation multiple for valuing firms in the South East Asian market.

Acknowledgments

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

References

Alford, A. (1992). The effect of the set of comparable firms on the accuracy of the price-earnings valuation method. *Journal of Accounting Research, 30*(1), 94-108.

Baker, M., & Ruback, R. (1999). Estimating industry multiple. Working paper. Harvard University.

Beaver, W., & Dale, M. (1978). What determines price-earnings ratios? *Financial Analyst Journal*, *34*(4), 65-76.

Boatsman, J., & Baskin, E. (1981). Asset valuation with incomplete markets. *The Accounting Review, 56*(1), 38-53.

Brealey, R. (1969). .*An Introduction to Risk and*

Return From Common Stocks. Cambridge: MIT Press.

Cragg, J., & Malkiel, B. (1968). The consensus and accuracy of some predictions of the growth of corporate earnings. *Journal of Finance, 23*(1), 67-84.

Copeland, T., Koller, T., & Murrin, J. (1994). *Valuation: Measuring and Managing the Value of Companies*. NY: John Wiley & Sons.

Damodaran, A. (1996). *Damodaran on Valuation: Security Analysis for Investment and Corporate Finance.* NY: John Wiley & Sons.

Huang, Y., Tsai, C., & Chen, C. (2007). Expected P/E, residual P/E, and stock return reversal: Time-varying fundamentals or investor overreaction?. *International Journal of Business and Economics, 6*(1), 11-28.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning With Application in R*. NY: Springer.

Kaplan, S., & Ruback, R. (1995). The valuation of cash flow forecasts: An empirical analysis. *Journal of Finance, 50*(4), 1059-1093.

Kasper, L. (1997). *Business Valuation: Advanced Topics*. CT: Quorum Books.

LeClair, M. (1990). Valuing the closely held corporations: The validity and performance of established valuation procedures. *Accounting Horizons, 4*(3), 31-42.

Liu, J., Nissim, D., & Thomas, J. (2002a). Equity valuation using multiples. *Journal of Accounting Research, 40*(1), 135-172.

Liu, J., Nissim, D., & Thomas, J. (2002b). International equity valuation using multiples. Working paper. UCLA Anderson School of Management.

Liu, J., Nissim, D., & Thomas, J. (2007). Is cash flow king in valuations?. *Financial Analyst Journal, 63*(2), 56-65.

Healey, P., Palepu, K., & Bernard, V. (2000). *Business Analysis and Valuation*. Ohio: South-Western College Publishing.

Penman, S. (1997). Combining earnings and book value in equity valuation. *Contemporary Accounting Research, 15*(3), 291-324.

Pereira, J., Basto M., & Ferreira da'Silva, A. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance, 39*, 634-641.

Pinto, J., Henry E., Thomas, R., & Stowe, J. (2016). *Equity Asset Valuation, CFA Institute Investment Series*. NJ: John Wiley & Sons.

Sehgal.S., & Pandey, A. (2010). Equity valuation using price multiples: Evidence from India. *Asian Academy of Management Journal of Accounting and Finance, 6*(1), 89-108.

Tasker, S. (1998). Industry preferred multiples in acquisition valuation. Working paper, Cornell University.