Stock Return Predictability by Bayesian Model Averaging: Evidence from Stock Exchange of Thailand

Kmonwan Chairakwattana
Sarayut Nathaphan

Thammasat Business School

This research paper examines the predictability power on future stock returns by employing the concept of Bayesian Model Averaging (BMA). The sample focuses on Stock Exchange of Thailand (SET) over 2001-2011. Predictors for return predictability contain financial information which are dividend yield, Book-to-Market, Earning yield, Default risk premium, Monthly rate of three-month Treasury bill, Term premium, Monthly inflation rate and Term spread. This paper also explores the predictability power over financial crisis, sub-period over 2008-2009. In addition, this paper compares expected returns from two models between BMA and traditional regression (Fama and Macbeth two steps procedure). Results indicated that BMA approach outperforms the traditional regression model.

Keywords: Stock Return, Investment Decision, Bayesian Model Averaging, Portfolio, Asset Pricing, Model Uncertainty

I. Introduction

The predictability in stock returns is an interesting issue by both practitioners and academicians. Many predictor variables have been employed in those literatures including dividend-price ratio, earnings-price ratio, firm size, interest rate and also other economic and fundamental variables. However, those empirical results did not explicitly indicate which variables should be included as predictors in predictive regression and there are concerned about data over fitting. In particular, Bossaerts and Hillion (1999) investigated predictability power on stock returns by separating data to in-sample and out-of-sample. They confirmed in-sample return predictability but failed to confirm out-of-sample return predictability. Their empirical results suggested that predictive variables used in regression should be treated as time varying factors. Historical data should be separated to prior and posterior distribution in order to reflect past information and expectations. Posterior probability of data should be conditional on information variables.

Recently, new methodology is employed for return predictability namely Bayesian Model Averaging (BMA). This model takes into account the conditional probability of posterior data. Most of previous empirical results regarding predictability on stock returns focused on developed market. For emerging market, there were a few studies about model uncertainty and none had applied Bayesian model to predictive regression. Therefore, it leads to the research question that "As one of the emerging capital market members, does Thailand stock market have return predictability power by applying Bayesian methodology?"

Many studies related to predictability of stock returns concluded that the predictability of stock returns mainly due to data on financial statements and macroeconomic variables. In accordance with asset pricing model, the expected rate of returns sensitive to beta coefficients which changes in the state of the economy (Ferson and Harvey, 1991). From the well-known asset pricing model, CAPM-APT model (Sharpe, 1964, Lintner, 1965 and Black, 1972) explained return correlated with only market risk factor.

Four major classifications on related variables used in return predictability are macro economic factors, size effect, market expectation as stated in book-to-market ratio, and behavioral factors. The first classification is macro economic factor, Pesaran and Timmermann (1995) investigated economic significance on return forecasts and stated that related economic variables were short and long interest rates, dividend yields, industrial production, company earnings, liquidity measures, and the inflation rate. These variables suggested by the variation between stock returns and business cycle (Prime, (1946), Dowrie and Fuller (1950), Rose (1960)). They studied US stocks during 1954 and 1992 and found that dividend yield was statistically significant in predicting stock returns. This result supported the previous studies by Campbell (1987) and Fama and French (1989).

The second classification is size effect pioneered by Banz (1981), he was the first one who suggested the relation between returns and the market value of common equity, normally this referred to as the size effect. His finding concluded that firm size had negatively relation with stock returns. The third classification is market expectation as shown in book-to-market Ratio. Fama and French (1992) suggested the relation between firm size and book-to-market ratios based on data from US stock exchange, excluding financial firms, during period 1963 to 1992. They concluded that three factors which are market, firm size, and book-to-market ratio had significant explanatory power to earnings. The fourth classification related to behavioral finance and phenomena. These variables are found important in previous studies including winners-minus-losers, value premium, January effect and also size premium. They are mostly notable as economy-wide factors in asset pricing models, supported by Merton (1973) and Campbell (1996). Furthermore, there are studies by Liew and Vassalou (2000) revealed that size premium and value premium are useful in predicting economic growth.

This paper focuses on Thailand stock market by applying the important variables found in the previous studies as well as those popular business cycle variables such as dividend yield, book-to-market, earning yield, interest rate on treasury bill, and also inflation rate, to investigate the predictability power of variables on stock returns. In addition, comparison between Bayesian Model Averaging (BMA) and other predictive regression models based on out-of-sample data is performed in order to suggest which model has better statistical predictive power.

II. Traditional Models for stock returns predictability

Most of empirical results regarding predictability on stock returns focused on traditional method. They started with regression model between excess return and predictor variables in lagged value. Later papers denoted the equation for predictor variables as stochastic process.

$$r_t = \alpha + \beta x_{t-1} + u_t \qquad (1)$$

$$x_t = \gamma + \rho x_{t-1} + e_t \qquad (2)$$

Where: r_t = excess return in period t
x_t = predictor variables in period t

John and Yogo (2006) employed conventional test in line with the regression above based on maximum likelihood ratio of the two equations and then testing for conventional t-test. Cooper et al. (2003) employed cross sectional regression similar to that of Fama and Macbeth (1973) aimed at determining significant predictor variables explaining excess return behavior in a multiple regression framework. Cross sectional regressions on lagged predictor variables were documented.

Rapach and Wohar (2005) investigated predictability of stock returns for both in-sample and out-of-sample. The methodology they used to assess in-sample predictability was t-statistic corresponding to the slope coefficient in a predictive regression model. However, the predictive regression model still does not impress for predictability of out-sample result. Many empirical results revealed lack of power to forecast the error for out-of-sample. There is some empirical result suggested for model selection to close the gap. Bossaerts and Hillion (1999), suggested statistical model selection criteria by using several criteria such as adjusted R^2, AIC, BIC, PIC and FIC. However, they discovered that event the best prediction model, it still lack of power to forecast out-of-sample period.

In addition, the traditional methodology does not take model uncertainty issue into account. Most of them ignore the conditional on data to impact changing in prior informative. This is maybe the reason for why predictive regression model lack of forecasting power for out-of-sample.

As Bossaerts and Hillion (1999), they found evidence in an international dataset by using linear models with varying numbers of predictors. They found that the out-of-sample forecasting power of retained prediction models was zero. In addition, they also found that no linear model generates any out-of-sample forecasting power. Indeed, the models were chosen with criteria that pick only the “best”. After that, there are some empirical results which taking such issue into consideration. They proposed the new methodology for predictability of stock returns in order to solve the issue for out-of-sample lacking of predictability power. Cremers (2001) investigated the ability of various variables to explain conditional expected returns. He found the important of conditional variables and a posterior probability of data is more important than the prior one. In addition, out-of-sample results for the Bayesian averages model showed improved forecast than the classical statistical model.

Avramov (2000) applied Bayesian weighted predictive distribution for return predictability model. Objective of Avramov was to indicate which predictor variables were the most likely or had the highest probability to be included in the predictive model. Based on out-of-sample data, by comparing forecast errors of Bayesian Model Averaging with that of other individual model selected by several criteria in accordance with Bossaerts and Hillion (1999), the result suggested that Bayesian has less forecasting error than other individual models.

Nathaphan and Chunhachinda (2010) explored optimal portfolio selection by using emerging market data as sample. The measurement of the better portfolio strategy was indicated by the higher expected utility, sharpe's ratio, and lower difference between ex-ante and ex-post average values from various strategy, including traditional mean variance (EV), Adjusted Beta, Resampled Efficient Frontier, Capital Asset Pricing Model (CAPM), Single Index Model, and Bayesian Single Index Model. Results stated that Bayesian Single Index Model or Bayesian portfolio performed best on an ex-ante and expost basis.

The existing literature provides a comprehensive examination of the predictive variables by using cross sectional regression and Bayesian Model Averaging and also found the evidence that Bayesian portfolio give the better result by comparing with other strategies. However, most of empirical results focused on developed market such as US and Japan. It has some room for emerging market such as Thailand to investigate the result regarding to predictive variable. In addition, Bayesian methodology is rarely use for emerging market data. Therefore, it has more advantage to investigate and contribute more evidence into financial industry for further studies regarding to Bayesian and emerging market in case of predictive on stock returns.

The Bayesian weighted predictive distribution

This paper would like to find out which predictive variables have explanatory power in predictive regressions by taking model uncertainty into consideration; therefore, this paper employs Bayesian framework as methodology to figure out the empirical result. By following Avramov (2002) which employed Bayesian Model Averaging in order to determine which predictive variables are significant in by applying this model.

The linear predictive regression should be employed when investors would like to predict future rate of returns on equity portfolios as follow:

$$r_t' = x_{j,t-1}' B_j + \varepsilon_t' \qquad (3)$$

When M explanatory variables are suspected relevant, there are 2^M competing regression specifications. In equation (3), $x_{j,t-1}'$ is a model-unique subset for each explanatory variable, ε_t' has been assumed as normally distribution with conditional mean zero and variance-covariance matrix $\in_j$

Bayesian model averaging computes posterior probabilities for the collection of all 2^M models. And then use probabilities as weights on the individual models to obtain the forecasting model which summarizes the dynamics of future stock returns.

Based on the conditional probability, the posterior probability of model j (denoted M_j) is given by

$$P(M_j|D) = \frac{P(D|M_j)P(M_j)}{\sum_{i=1}^{2^M} P(D|M_j)P(M_j)} \qquad (4)$$

Where: D stand for the data $P(M_j)$ is the prior distribution of M_j and $P(D|M_j)$ is the corresponding marginal likelihood.

After that, the process to find out marginal likelihood of the model j has been employed. From the posterior probability, it means that model j – which means each predictive variable, probability conditional on the data. The marginal likelihood of M_j is given by

$$P(D|M_j) = \frac{L(\in_j, B_j, D, M_j) P(\in_j, B_j|M_j)}{P(\in_j, B_j|D, M_j)} \qquad (5)$$

Where: $L(\in_j, B_j, D, M_j)$ is the likelihood function to M_j and $P(\in_j, B_j|M_j)$ and $P(\in_j, B_j|D, M_j)$ are the joint prior and posterior distribution of the model specific parameters, respectively.

From model j, we assume that data containing in that model is stochastic process denoted by

$$y_{j,t}' = x_{j,t-1}' \theta_j + u_{j,t}' \qquad (6)$$

Where: $u_{j,t}'$ is the error term of predictive variable, assumed as independent and identically distributed normal variables which has mean = 0 and variance denoted by φ_j. From equation (3) and (6), both of them have error term, equation (3) stands for ε_t' and equation (6) stands for $u_{j,t}'$

After that, Vector autoregressive (VAR) has been employed to combined equation (3) and (6) altogether.

$$Z'_{j,t} = a'_j + Z'_{j,t-1}A_j + n'_{j,t} \qquad (7)$$

Follow Avramov (2002), the general form of Bayesian weighted predictive distribution of cumulative excess continuously compounded returns average over the model that takes uncertainty of coefficient of predictive variable (θ_j) and variance of error term (φ_j)

$$P(R_T|D) = \sum_{j=1}^{2^M} P(M_j\,|D) \int_{\theta_j,\varphi_j} P(\theta_j,\varphi_j\big|M_j,D\big)P\big(R_T\big|M_j,\theta_j,\varphi_j,D\big)d\theta_j,d\varphi_j \qquad (8)$$

From prior distribution, denoted as $P(M_j)$, and then I would like to find posterior distribution conditional on data in particular model. I find the particular which does not overfitting the data by selection the model is averaging maximized on likelihood function $P(D|M)$. After that, we use such model in our general form, as equation (8).

In equation (8), it applied Bayesian Model Averaging (BMA). The first part which is $\sum_{j=1}^{2^M} P(M_j\,|D)$- the sum of posterior distribution- particular model is averaging by marginal likelihood as in equation (5). In this paper, I employed 14 variables which mean I have 14 candidate models. The second part which is $\int_{\theta_j,\varphi_j} P(\theta_j,\varphi_j\big|M_j,D\big)P\big(R_T\big|M_j,\theta_j,\varphi_j,D\big)d\theta_j,d\varphi_j$ – means the posterior distribution of specific parameter given data on that particular model. In conclusion, the general form came from three steps. First of all, drawing from the discrete distribution model, model M_j is drawn for probability P(M_j|D). It means model M_j has conditional probability on data (D). Second, model specific parameters θ_j,φ_j are drawn from their joint posterior distribution. The last one, given θ_j,φ_j, they are vectors of cumulative excess continuously compounded returns is drawn from the distribution of future stock returns conditional on the model which are the specific parameter (θ_j,φ_j) of predictive variables and sample data.

The predictive probability of seeing particular value of a new observation will vary depending on the parameters of the distribution of the observation. Although, the exact value of the parameters are hardly to specify, but the posterior distribution can be identify over them in accordance with believe the parameters to be, given the data already seen. If compare this to the traditional frequents statistics, where a single estimate of parameters e.g. maximum likelihood estimate, would be computed. This is equivalent to averaging over posterior distribution with no variance. The result is weighted too strongly towards the mode of posterior, and takes no account of other possible values.

After obtained the posterior probabilities, this paper propose three statistics to investigate which explanatory variables have the most power in predictive regressions.

1. Cumulative posterior probabilities of the predictive variables. This statistic will quantify the probabilities that each of the predictive variables appears in the weighted forecasting model.
2. Posterior t ratio obtained by dividing the posterior mean of each of the slope coefficients in the weighted model by its corresponding posterior standard error. This statistics will measure how much the predictive variable have statistically significant in the predictive regression. Based on Avramov (2002), the paper suggested that the greater model uncertainty the smaller the posterior t-ratio or the less likely is the predictor to be statistically significant.
3. Posterior-odds ratio obtained by dividing the sum of posterior probabilities (which retains at least one predictor) by the posterior probability of the iid model. This statistic will be tested for portfolio efficiency which suggested by Shanken (1987).

After the testing of three statistics mentioned above, currently the result will present which predictor variables have the most explanatory power to predictive regression. And then, the top three will be selected to perform the further step.

II. Data

Data used for this study is monthly observation on stock returns and information variables over January 2001 through December 2011. There are 439 listed companies listed on Stock Exchange of

Thailand (SET) as at 31 December 2011. The sample excluded listed leasehold right and property fund and the companies which are under restructuring process. In sample data ranges from January 2001 to December 2011 and out-of-sample ranges from January 2012 to December 2012 with the sample size of 494 companies as at 31 December 2012.
The investment universe consists of the six Fama and French (1993) portfolios, formed as two sizes which are small (S) and big (B) and three book-to-markets which are low (L), medium (M) and high (H). Each of the 2^M competing models consist of a unique subset of the following M = 8 information variables (taking one lag)

Information variables	Definition	Source
1. Dividend yield (Div)	The most recently announced net dividend, annualized based on the Dividend Frequency, then divided by the current market price.	SETSMART
2. Book-to-Market (BM)	Measure of the relative value of a company to its market value. Calculated as Book value divided by market capitalization.	SETSMART
3. Earning yield (EY)	Earnings per share expressed as a percentage of share price.	SETSMART
4. Default risk premium (Def)	The difference between return on long-term corporate bonds (proxy by WACC cost of debt) and return on long-term government bond (proxy by 10 year government bonds)	SETSMART and ThaiBMA
5. Monthly rate of a three-month Treasury bill (Tbill)		ThaiBMA
6. Term premium (Term_p)	The difference between the monthly return on long-term government bond (proxy by 10 year government bond) and the one-month Treasury bill rate	ThaiBMA
7. Monthly inflation rate (Inf)		BOT
8. Term spread (Term_s)	The difference in annualized yield of ten-year and one-year Treasuries.	ThaiBMA

The predictor variables have been selected based on the important found in the previous studies and also the popularity of such variables in business cycle. From the traditional regression, it presents that selected predictive variables are statistically significant correlated with gross returns as shown in Table I.

Table I

Multiple regression of monthly gross return on predictive variables

Portfolio	Predictive variables							
	Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
Individual	0.00	-0.01	0.00	-0.01	0.00	-0.01	-0.02	0.01
	-1.07	-10.45	4.33	-10.40	0.61	-5.18	-14.45	5.75
	0.00	-0.01	0.10	-0.01	-0.02	0.00	-0.01	0.01
SL	0.00	0.00	0.00	0.00	0.00	-0.03	-0.02	0.02
	0.04	0.65	-0.36	-0.76	0.72	-1.95	-1.99	2.42
	0.00	0.00	-0.17	-0.01	0.00	0.01	0.00	0.02
SM	0.00	-0.04	0.00	-0.02	0.00	-0.02	-0.02	0.00

	1.66	-5.67	4.32	-6.83	1.02	-2.89	-4.76	1.63
	0.01	0.00	0.33	-0.01	-0.01	0.00	-0.01	0.00
SH	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.02	0.01
	2.59	-6.80	-0.34	-3.88	0.13	-3.36	-7.33	4.12
	0.03	-0.01	0.09	-0.01	-0.02	0.01	-0.01	0.01
BL	-0.01	-0.01	0.00	-0.01	0.00	-0.01	-0.02	0.00
	-8.63	-1.85	4.28	-2.65	-1.17	-0.93	-5.29	1.20
	-0.06	0.00	0.25	-0.01	-0.02	0.00	-0.01	0.01
BM	0.00	-0.04	0.00	-0.02	0.00	-0.01	-0.02	0.01
	-3.57	-7.99	2.31	-10.21	0.24	-3.67	-10.65	3.67
	-0.03	0.00	0.02	-0.02	-0.02	0.01	-0.01	0.01
BH	0.00	-0.02	0.00	-0.01	0.00	-0.02	-0.03	0.01
	-2.06	-8.53	-0.41	-6.77	-0.33	-3.39	-8.59	4.13
	-0.02	-0.01	0.03	-0.02	-0.02	0.01	-0.01	0.02

The table exhibits slope coefficients (first row) and their corresponding t-ratios (second row) obtained from regressing gross return on predictive variables on each of the size book-to-market portfolios. The third row is covariance between gross returns and predictive variables. The predictive variables include; dividend yield (Div), book-to-market (BM), earning yield (EY), The difference between return on long-term corporate bonds (proxy by WACC cost of debt) and return on long-term government bond (proxy by 10 year government bonds) (Def), the monthly rate of a three-month Treasury bill (Tbill), The difference between the monthly return on long-term government bond (proxy by 10 year government bond) and the one-month Treasury bill rate (Term_p), Inflation rate (Inf) ; and The difference in annualized yield of ten-year and one-year Treasuries (Term_s)

III. Empirical Results

A) Bayesian Model Averaging: Stock return predictability (During 2001-2011)

In accordance with concept of Bayesian Model Averaging (BMA), this paper considered of all linear data-generating process of 8 predictive variables. From all 8 variables, it will be come up with the comparison of 2^8 = 256 models. BMA will compute the marginal likelihood for every model and then weights the marginal likelihood by the model prior probability and normalizes the result to obtain the model posterior probability. This method is employed by the BMS Package. This package employs standard Bayesian normal-conjugate linear model, usually known as the unit information prior (UIP) and then use a Markov Chain Monte Carlo to come up with the most important model.

Table II

Posterior probabilities of forecasting models based on a prior sample weighted against predictability

Portfolio	Predictive variables							
	Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
Individual	0.02	1.00	0.98	1.00	0.02	1.00	1.00	1.00
	0	**1**	0	**1**	0	**1**	**1**	**1**
SL								

	0.99 **1**	0.05 0	0.07 0	0.02 0	0.02 0	0.36 0	0.04 0	0.63 **1**
SM	0.08 0	1.00 **1**	1.00 **1**	1.00 **1**	0.03 0	0.87 0	1.00 **1**	0.10 0
SH	0.30 0	1.00 **1**	0.02 0	0.98 0	0.02 0	0.52 0	1.00 **1**	0.64 0
BL	1.00 **1**	0.09 0	0.90 0	0.98 0	0.06 0	0.02 0	1.00 **1**	0.03 0
BM	0.48 0	1.00 **1**	0.07 0	1.00 **1**	0.01 0	0.11 0	1.00 **1**	0.11 0
BH	0.20 0	1.00 **1**	0.04 0	1.00 **1**	0.03 0	0.62 0	1.00 **1**	0.74 0

The table reports results from BMA. The first row displays cumulative posterior probabilities. The second row denotes the highest posterior probabilities by assigning one for the highest probabilities among other variables. It supposes that variables which have been assigned one should include as predictive variables for calculation expected return of out sample data. The stock universe consists of six portfolios identified by two letters which are size (S,B) and book-to-market (L,M,H) during period 2001 to 2011. Following are the predictive variables set: dividend yield (Div), book-to-market (BM), earning yield (EY), The difference between return on long-term corporate bonds (proxy by WACC cost of debt) and return on long-term government bond (proxy by 10 year government bonds) (Def), the monthly rate of a three-month Treasury bill (Tbill), The difference between the monthly return on long-term government bond (proxy by 10 year government bond) and the one-month Treasury bill rate (Term_p), Inflation rate (Inf) ; and The difference in annualized yield of ten-year and one-year Treasuries (Term_s).

As the result in Table II, consider individual stock which is the stock universe without any classification. The evidence supporting return predictability, as many of the information variables have highly posterior probabilities. It means each of predictive variables have high probability to appear in the weighted forecasting model. The highest probability predictors are book-to-market, default risk premium and inflation rate. The inflation rate is capability in predicting both small-cap and large-cap stocks. These predictive variables reflects Thai stock market environment which has been normally nominated by international financial institution, mutual fund and also foreign investors. These kinds of investors mostly make decision to invest in accordance with economic factors of such country, including inflation rate.

Book-to-Market ratio is also prominent to predict return for both small-cap and large-cap stocks, however, the capability are focused only value stocks. It poorly predicts growth stocks (Fama and French, 1993). Default risk premium better predicts large-cap stocks than small-cap stocks, especially for value stocks. This variable is proxied to risk premium that borrower has to pay in additional from higher risk in the firm. Normally, small-cap stocks are higher risk which affect from other risk factors and credit rating. The investors in large-cap stocks have to pay attention to risk premium which each firm has to bear from cost of debt. Next, among the traditional market ratios which are dividend yield, book-to-market and earnings yield. Except book-to-market, dividend yield is prominent in predicting growth stocks regardless of their size classification.

Focusing on default risk premium, it is proxied by cost of debt of each firm minus risk free rate. The risk free rate in this research proxied by government bond 10 years (Avramov, 2001) and found the predictability power in favor of large-cap stocks. In this section, the paper tries to come up with whether short-term and long-term yield impact over stock returns. When apply government bond 1 year as proxy for risk free rate and then calculated for default risk premium. We found that default risk premium has less predictability powers by comparing with proxied by 10 year government bond. (Table III). From the evidence, yield curve on term structure of interest has power on stock returns. In accordance with Avramov (2001), shifts in interest rates and economic conditions are related to the degree of risk aversion embedded in stock pricing. If pay attention to maturity, the long-term maturity on government bond suppose to reflect the economic condition better than short-term government bond. Stock Exchange of Thailand (SET) is quite sensitive to economic condition; therefore, default risk premium which proxied by 10-year government bond is in better predictors than proxied by 1-year government bond.

Table III

Posterior probabilities after changing default risk premium calculation

Portfolio	Predictive variables								
	Div	BM	EY	Def (1 yr)	Def (10yr)	Tbill	Term_p	Inf	Term_s
Individual	0.02	1.00	0.96	0.01	1.00	1.00	0.89	1.00	1.00
SL	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.01
SM	0.04	0.99	0.99	0.01	1.00	0.51	0.02	1.00	0.06
SH	0.17	1.00	0.01	0.01	0.98	0.04	0.23	1.00	1.00
BL	1.00	0.07	0.91	0.04	0.98	0.76	0.10	1.00	0.30
BM	0.70	1.00	0.19	0.07	1.00	1.00	0.17	1.00	1.00
BH	0.16	1.00	0.03	0.04	1.00	0.87	0.40	1.00	1.00

Table III represents posterior probabilities for data during 2001-2001 by changing calculation for default risk premium (Def) variables.

Table IV

Slope coefficients in the weighted model and posterior t-ratios

Portfolio	Predictive variables							
	Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
Individual	0.000	-0.007	0.000	-0.011	0.000	-0.012	-0.020	0.006
	-0.115	-10.493	3.707	-11.352	0.065	-5.004	-16.707	5.502
SL	-0.004	0.000	0.000	0.000	0.000	0.005	0.000	0.004
	-4.356	-0.194	0.239	-0.091	-0.120	0.704	-0.168	1.206
SM	0.000	-0.039	0.000	-0.019	0.000	-0.009	-0.015	0.000
	0.251	-5.217	4.678	-5.849	0.126	-1.758	-4.466	0.099

SH	0.000	-0.014	0.000	-0.010	0.000	-0.010	-0.022	0.006
	0.594	-6.589	-0.018	-3.544	-0.017	-0.942	-6.020	1.097
BL	-0.006	-0.001	0.000	-0.010	0.000	0.000	-0.020	0.000
	-8.868	-0.272	2.282	-3.925	-0.203	0.055	-6.746	0.142
BM	0.000	-0.036	0.000	-0.016	0.000	-0.001	-0.021	0.001
	-0.874	-7.710	0.251	-14.100	0.004	-0.329	-12.586	0.328
BH	0.000	-0.019	0.000	-0.016	0.000	-0.011	-0.025	0.006
	-0.441	-8.574	-0.119	-6.984	-0.077	-1.143	-7.505	1.327

The first row represents posterior means of slope coefficients obtained by averaging slope estimates across models. The second row denotes t-ratios to account of model uncertainty which obtained by dividing the posterior mean of each of the slope coefficients by its posterior standard error.

Table IV presents posterior means of slope coefficients in the weighted model and t-ratios which obtained by dividing the posterior mean by its posterior standard error. The t-ratios is the measurement of model uncertainty regarding to the apparent predictive power of variables. The greater t-ratios mean less model uncertainty and it can interpret that the higher posterior probability variables are more likely to be statistically significant.

As suspected, the posterior probabilities are related to the posterior t-ratios. Both small-cap and large-cap portfolios, high posterior probabilities for the book-to-market, default risk premium and inflation rate (Table II) are followed by higher values of posterior t-ratios (Table IV), especially for value stocks (middle and high book-to-market ratio). Focusing on dividend yield, smaller posterior probabilities in both small-cap and large-cap portfolios are also followed by smaller posterior t-ratios. The greater posterior t-ratios mean that book-to-market, default risk premium and inflation rate are more likely the predictors to be statistically significant.

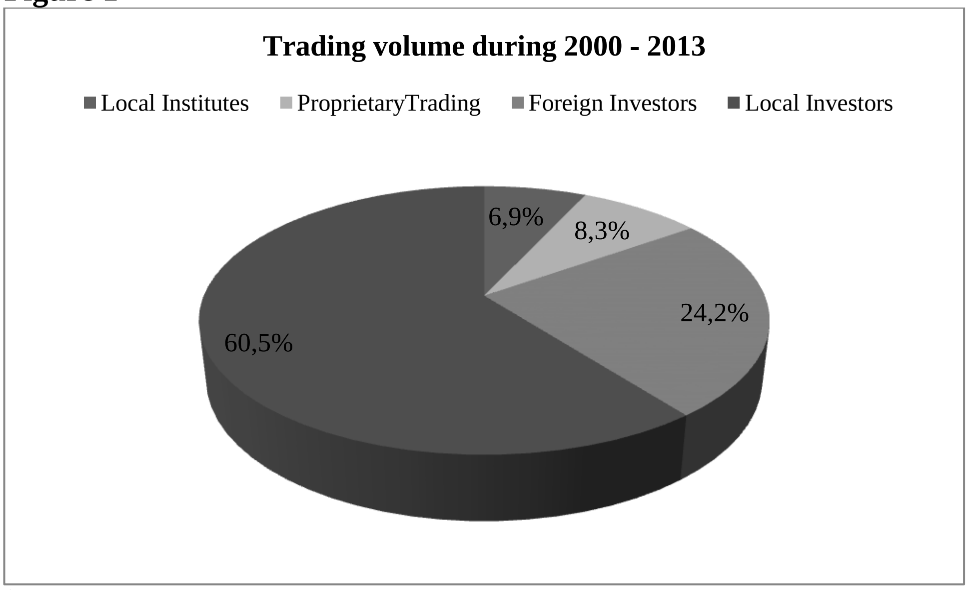
The third statistic is posterior-odds ratio of predictability versus no predictability. This measurement reveals the portfolio efficiency to evaluate which portfolios are in favor of predictability. Posterior-odds for the six size book-to-market portfolios are 2.17 (SL), 5.08 (SM), 4.47 (SH), 4.07 (BL), 3.79 (BM) and 4.63 (BH). These figures are as evidence for cross sectional differences in predictability. The evidence in favor of predictability is the strongest for small value stocks (SM), the weakest for small-growth stocks (SH). Holding book-to-market fixed the favor of predictability power contributed to large-cap stocks relative to small-cap stocks (4.07 versus 2.17 for low book-to-market stocks and 4.63 versus 4.47 for high book-to-market stocks). Similarly, controlling for size, posterior-odds are higher for high-versus-low book-to-market stocks (4.47 versus 2.17 for small-cap stocks and 4.63 versus 4.07 for large-cap stocks). From three statistical from BMA approach, during normal economics circumstance, the predictability powers of variables are book-to-market, default risk premium and inflation rate. Out of traditional ratio such as dividend yield, earnings yield and book-to-market ratio which have already revealed the power by other researches, have relatively small posterior probabilities of being correlated with future returns, except book-to-market ratio.

This research paper comes up with three predictor variables which are book-to-market, default risk premium and inflation rate. In addition, the evidence also reveals that large-cap value stocks are in favor of predictability power. Unlike that of Avramov (2001), his results found that the robust predictors are term premium and market premium and more powerful over small-cap value stocks. The evidence can be explained by focusing on sample used to examine the predictability on stock returns. Avramov (2001) examine overall US stock market during period April 1953 through

December 1998, while this paper focus on Thailand stock market during period January 2001 through December 2011. The difference environment of stock market between USA and Thailand is explicit among investor's perception. US stock market is treated as developed and efficient market. Information has been widespread to all investors. The term premium captures exposure related shift in interest rates and economic conditions. This variable is related to the degree of risk aversion embedded in pricing. Market premium predictability power over small-cap stocks means current return on small-cap stocks are correlated with past return of large-cap stocks. In addition, the evidence also reveals more powerful of predictors on small-cap stocks than large-cap stocks. The predictors confirmed the efficient market in US stocks. They have predictability on future stock returns over small-cap stocks which do not have any dominated by big players in the market. Unlike Stock Exchange of Thailand (SET), the evidence examined over such sample revealed that the powerful predictors are book-to-market, default risk premium and inflation rate in favor of large-cap value stocks. In accordance with Chordia and Swaminathan (2000), find that high volume stocks (mostly large-cap stocks) respond rapidly whereas low volume stocks (mostly small-cap stocks) respond slowly to marketwide information. Normally, Stock Exchange of Thailand (SET) is sensitive market to economics information. Because of emerging market and have been controlled by some influence of investors, this leads to more sensitive on external environment factors.

If deeply examined the details of investors in market, the evidence found that the influence group of investors who dominated market are foreign investors (Table V). Evidence in Table V presents correlation between trading value with stocks index and found that foreign investors and proprietary trading have positive relations with index, while local investors have negative relations. Although, when I performed further tested whether foreign investors dominate SET by using Granger Causality test, the evidence presented that they do not lead each other. But, we found the causality effect between foreign and local investors. It means local investors have been dominated by foreign investors to make decision for investment. Normally, these kinds of investors consider overall environment of invested countries, especially macro environment. Macroeconomics factors are reflected in inflation rate and default risk premium (in term of interest rate). Therefore, these variables are more powerful on predictability stock returns in Stock Exchange of Thailand (SET), especially in favor of large-cap value stocks.

Figure 1[1]



[1]Source: www.setsmart.com

Table V

Trading volume correlation with Stock Exchange of Thailand (SET) Index

Investor types	Correlation with Index
Local Institutes	(0.0416)
Proprietary Trading	0.0709
Foreign Investors	0.3395
Local Investors	(0.4299)

Table V denoted trading volume correlation between trading volume of each investor types and Stock Exchange of Thailand (SET) Index. By using data during 2000-2013, the evidence revealed that foreign investors have positive correlation with SET; therefore, this type of investor can dominated movement trend of market.

B) Bayesian Model Averaging: Stock return predictability (Financial crisis 2008 to 2009)

The result above presents the stock return predictability based on BMA model by using sample during 2001 to 2011. Unfortunately, Thailand was in financial crisis during 2008 to 2009. This crisis significant affected to overall economics, including Stock Exchange of Thailand (SET). In order to confirm the existence of stock return predictability under financial crisis, this paper separated data under this financial crisis (during 2008 to 2009) to perform regression under BMA.

Table VI presents posterior probabilities for data under Thailand financial crisis (2008 to 2009). Focusing on small-cap portfolios, the highest posterior probabilities are default risk premium and inflation rate. As similar to large-cap portfolios, the highest posterior probabilities are also default risk premium, inflation rate and book-to-market; this is different from small-cap portfolios. By comparing with sample during normal circumstance, the predictor variables are quite the same. Book-to-market, default risk premium and inflation still are the highest posterior probabilities in both normal and crisis circumstances. High posterior probabilities for book-to-market, default risk premium and inflation rate are followed by higher values of posterior t-ratios for both portfolios.

Table VI

Posterior probabilities of forecasting models based on a prior sample weighted against predictability

Portfolio	Predictive variables							
	Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
Individual	0.14	1.00	0.93	1.00	0.95	0.99	1.00	1.00
	0	**1**	0	**1**	0	0	**1**	**1**
SL	0.17	0.06	0.03	0.70	0.15	0.02	0.06	0.02
	0	0	0	**1**	0	0	0	0
SM	0.12	0.58	0.78	1.00	0.03	0.05	1.00	0.10
	0	0	0	**1**	0	0	**1**	0
SH	0.02	1.00	0.03	0.92	0.07	0.11	0.99	0.45
	0	**1**	0	0	0	0	**1**	0
Portfolio	Predictive variables							

	Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
BL	0.23	0.31	0.13	0.03	1.00	0.77	0.06	0.99
	0	0	0	0	**1**	0	0	**1**
BM	0.98	1.00	0.07	1.00	0.30	0.25	1.00	0.73
	0	**1**	0	**1**	0	0	**1**	0
BH	1.00	1.00	0.11	1.00	0.09	0.33	1.00	0.90
	1	**1**	0	**1**	0	0	**1**	0

The table reports results from BMA. The first row displays cumulative posterior probabilities. The second row denotes the highest posterior probabilities by assigning one for the highest probabilities among other variables. It supposes that variables which have been assigned one should include as predictive variables for calculation expected return of out sample data. The stock universe consists of six portfolios identified by two letters which are size (S,B) and book-to-market (L,M,H) during the period 2008 to 2009.

Table VII

Slope coefficients in the weighted model and posterior t-ratios

Portfolio	Predictive variables							
	Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
Individual	0.000	0.010	0.000	0.019	0.013	0.025	0.036	0.017
	-0.307	-6.474	2.466	-5.628	-2.790	-4.051	-6.585	6.165
SL	0.000	0.001	0.000	-0.017	-0.003	0.000	-0.001	0.000
	0.405	0.220	0.136	-1.357	-0.379	-0.089	-0.211	0.010
SM	0.000	-0.024	0.000	-0.030	0.000	0.000	-0.046	0.001
	0.324	-1.031	1.595	-6.192	-0.009	0.089	-4.592	0.274
SH	0.000	-0.025	0.000	-0.020	-0.001	0.001	-0.055	0.004
	0.018	-5.202	-0.094	-2.429	-0.183	0.190	-4.587	0.796
BL	-0.001	-0.021	0.000	0.000	-0.035	-0.034	-0.001	0.023
	-0.485	-0.602	-0.340	-0.069	-6.347	-1.549	-0.196	2.417
BM	-0.003	-0.060	0.000	-0.027	-0.004	0.002	-0.066	0.005
	-3.372	-5.579	0.182	-4.698	-0.574	0.410	-6.748	1.385
BH	-0.006	-0.031	0.000	-0.033	0.000	-0.009	-0.107	0.013
	-4.724	-6.197	-0.227	-5.620	0.091	-0.503	-10.089	1.455

The first row represents posterior means of slope coefficients obtained by averaging slope estimates across models. The second row denotes t-ratios to account of model uncertainty which obtained by dividing the posterior mean of each of the slope coefficients by its posterior standard error. Stock universe is in six size book-to-market portfolios during 2008 to 2009.

Posterior-odds ratio is the last statistic figure which has been taken into account in order to evaluate the efficiency of portfolios in favor of predictability during financial crisis. As similar to normal circumstances, holding book-to-market, large-cap stocks has higher value relative to small-cap stocks (5.52 versus 2.20 for low book-to-market stocks and 9.43 versus 5.57 for high book-to-market stocks). Then, controlling size, posterior-odds has higher in high book-to-market stocks than low book-to-market stocks (5.57 versus 2.20 for small-cap stocks and 9.43 versus 5.52 for large-cap stocks). From the last statistic figure, large-cap value stocks portfolio is in favor on predictability power under financial crisis.

Book-to-market, default risk premium and inflation rate are predictability variables for both normal and financial crisis circumstances from this research paper's result. In addition, both of them are in favor of large-cap value stocks portfolio. As mentioned, large-cap stocks respond more rapidly than small-cap stocks in term of marketwide information. Therefore, it is not surprised that the powerful of predictors still be around large-cap stocks. However, the predictors which have predictability on future stock returns under financial crisis should be difference with under normal environment. Focusing on Stock Exchange of Thailand (SET), as mentioned earlier regarding to correlation between investor type and SET Index, Foreign investors are denominated SET Index in positive relations. This group of investor has quite systematically invests and carefully concern about macro economics as usual (As explained in section normal circumstances). Although the result is quite similar between normal and under financial crisis, this result of robustness still has difference issue. Focusing on posterior-odds ratio, average of this statistic compute from normal period is higher than under financial crisis. This evidence could be possible that BMA is in favor for portfolio efficiency during normal circumstance.

C) Bayesian Model Averaging: Out-of-sample performance

In a related study, Avramov (2001) confirmed the presence of predictability using BMA over out-of-sample by using several model selection criteria refer to Bossaerts and Hillion (1999). He revealed that the out-of-sample performance of the Bayesian approach is superior to that of model selection criteria. Our study focus on the properties of forecast errors generated by the Bayesian approach and the traditional return generating regression.

Our examination focuses on comparing mean squared error (MSE) between Bayesian approach and traditional regression (Fama Macbeth 2 steps procedure). MSE is calculated by using coefficient of each predictability variables for each portfolio.

Table VIII

Mean Squared Error

Portfolio	Approach				
	BMA	FMB	Difference	p-value (pair t-test)	p-value (non-parametric)
Individual	0.0371	0.0405	0.0034	0.1631	0.1143
SL	0.0381	N/A	N/A	N/A	N/A
SM	0.0564	0.0865	0.0301	0.0237	0.0124
SH	0.0369	0.4237	0.3868	0.0000	0.0356
BL	0.0286	0.0474	0.0188	0.0016	0.0023

BM	0.0377	0.0416	0.0039	0.1228	0.2225
BH	0.0262	0.0581	0.0318	0.0000	0.0005

Table VIII denoted Mean Squared Error (MSE) for both approaches. Mean squared error calculated by using coefficient obtained from each approach. Posterior mean obtained from BMA and traditional coefficients obtained from Fama Macbeth (2 steps procedure) are used to calculate expected return. Then, compare expected return with actual return obtained from out-of-sample (data in year 2012); use the difference to calculate MSE. The last column presents pair t-test for error of MSE in each portfolio. The result confirmed that most of them are significant statistically with 90% confidential.

The out-of-sample forecast errors of traditional return generating process, Fama and Macbeth (2 steps procedure) displays undesirable values. Forecasts are not efficient, though the coefficients in the regression are significant statistically. Focus on small-cap portfolio, the performance of forecasting future returns are inefficient. Especially for small-cap growth stocks portfolio, there are no any predictability power to expected return, while Bayesian approach has at least one predictor to forecast future returns. Similar to large-cap portfolio, coefficients obtained from traditional regression are still lack of power to forecast future returns. In contrast, the result shows that Bayesian Model Averaging has an impressive out-of-sample performance. Mean squared error for all portfolios are less than the traditional approach. Specifically, based on the monthly sample, mean squared errors generated by Bayesian model averaging range between 0.02 and 0.05, while Fama-Macbeth 2 steps procedure generates at range between 0.02 and 0.4. The wider of range means the more difference from actual return. Furthermore, this research paper examine the difference between MSE whether it has difference in significant statistical. Based on pair t-test, we found that the differences are significant statistical with 90% confidentiality. Furthermore, this paper employed non-parametric test because of small sample, the result confirms similar evidence as pair t-test. Therefore, Bayesian Model averaging has outperformed by comparing with traditional regression on out-of-sample in accordance with Avramov (2001).

Table IX
Coefficient used to calculate expected return for out-of-sample period (Year 2012)

Portfolio	Approach	Predictive variables							
		Div	BM	EY	Def	Tbill	Term_p	Inf	Term_s
Individual	BMA	n/a	-0.0069	0.0001	-0.0107	n/a	-0.0123	-0.0198	0.0059
	FMB	n/a	n/a	n/a	-0.0117	n/a	-0.0089	-0.0242	0.0033
SL	BMA	-0.0040	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	FMB	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SM	BMA	n/a	-0.0387	0.0003	-0.0188	n/a	-0.0093	-0.0148	n/a
	FMB	n/a	n/a	n/a	0.0497	n/a	n/a	n/a	-0.0266
SH	BMA	n/a	-0.0142	n/a	-0.0101	n/a	n/a	-0.0220	n/a
	FMB	n/a	-0.1717	n/a	n/a	n/a	n/a	-0.0925	n/a

BL	BMA	-0.0062	n/a	0.0002	-0.0103	n/a	n/a	-0.0198	n/a
	FMB	n/a	n/a	n/a	0.1386	n/a	n/a	n/a	n/a
BM	BMA	n/a	-0.0362	n/a	-0.0164	n/a	n/a	-0.0213	n/a
	FMB	n/a	-0.1994	n/a	n/a	n/a	n/a	n/a	n/a
BH	BMA	n/a	-0.0194	n/a	-0.0158	n/a	n/a	-0.0252	n/a
	FMB	n/a	-0.1398	n/a	n/a	n/a	n/a	n/a	-0.0170

Table IX presents coefficient used to calculate expected return for each approach. BMA stand for Bayesian Model Averaging, while FMB stands for Fama Macbeth 2 steps procedure. This paper focuses on only predictability variables on significant statistically which obtained from in sample. Out-of-sample is data in period 2012.

IV. Conclusion

This study focuses on stock return predictability based on Bayesian Model Averaging over period 2001-2011 and takes into account the impact of Asian Financial Crisis during 2008-2009. Out-of-sample in period, Jan 2012 to Dec 2012 is used to comparing the efficiency of forecasting future returns between Bayesian Model Averaging and Traditional regression approach. In the context of predictive regressions, the Bayesian methodology is attractive. By using posterior probabilities to a wide set of competing return-generating models, which are difference variables; it uses the probabilities as weights on the individual models to obtain a composite weighted model. The posterior probabilities represent the quantity of probabilities that each of the predictive variables appears in the weighted forecasting model. From the posterior probabilities, this research paper found that book-to-market, default risk premium and inflation rate are useful predictors of future stock return, especially in favor of large-cap stocks.

During financial crisis period 2008-2009, the predictability power still be found over such period. Especially, default risk premium is important factors during crisis. It is attractive for both small-cap and large-cap portfolio. Therefore, during financial crisis, consideration of investors is similar to normal period. Normally, foreign investors who are influence investors over SET have systematically and carefully consideration regarding to investment for both normal and crisis situations.

Focus on out-of-sample during 2012, this research paper found Bayesian Model averaging outperforms the traditional regression model; Fama Macbeth two steps procedure (FMB). By using significant statistically coefficients to compute expected returns of out-of-sample and then calculate for Mean Squared Error (MSE). The differences between MSE from two approaches are significant statistical difference with 90% confidentiality after applying pair t-test. The evidence support BMA is more attractive than FMB.

In conclusion, Bayesian Model Averaging (BMA) can be used to examine return predictability in the context of Emerging market, especially Thailand. This model is efficient in both normal and crisis circumstances and also still have predictability power over out-of-sample.

References

Avramov, 2002, Stock Return Predictability and model uncertainty, *Journal of Financial Economics* 64, 423-458

Peter Bossaerts, Pierre Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What do we learn?, *The Review of Financial Studies* 12, 405-428

Fama, French, 2004, The Capital Asset Pricing Model: Theory and Evidence, *Journal of Economic Perspective* 18, 25-46

Fama, French, 1995, Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance* 1, 131-155

Cremers, K.J.M., 2002, Stock Return Predictability: A Bayesian model selection perspective, *Review of Financial Studies* 15, 1223-1249

Campbell, J.Y. and R.J. Shiller, 1988, Stock prices, Earnings, and expected dividends, *Journal of Finance* 43, 661-676

Campbell, J.Y., 1987, Stock returns and term structure, *Journal of Financial Economics* 18, 373-399

Nathaphan, and Chunhachinda, 2010, Estimation Risk Modeling in Optimal Portfolio Selection: An Empirical Study from Emerging Markets, *Economics Research International* 2010, 10 pages

Donald B. Keim and Robert F. Stambaugh, 1986, Predicting Returns in the Stock and Bond Markets, *Journal of Financial Economics* 17, 357-390

Fama E., Schwert, G.W., 1977, Asset Returns and Inflation, Journal of Financial Economics 5, 115-146.