Comparing Regularization Techniques for Thai Exports Growth Forecasting Model

Student Name: [redacted]

Supervisor Name: [redacted]

Submitted as part of the degree of M.Sc. MISCADA to the

Board of Examiners in the Department of Computer Science, Durham University.

Abstract –

Context/Background – Export forecasts have significant economic implications for business policy-making decisions. By far, economists and practical forecasters employ the time-series models. While preserving parsimony, these methodologies cannot take in the information from the broader sets of potential predictors such as country-level trading partner GDP and the movement of bilateral exchange rates. Thus, I examine the regularization techniques to construct the forecasting model based on these potential predictors.

Aims – This dissertation aims to exploit several regularization techniques to specify the forecasting model for Thai export growth. I aim to compare the forecasting performance of several models constructed based on different regularization techniques and to contrasts these models with the traditional time-series model and factor model.

Method – I conduct a variable shrinkage from the rich dataset of potential predictors using Lasso, Adaptive Lasso, Elastic Net, and Adaptive Elastic Net method. The shrinkage model is estimated using mainly the Least Square model, while Bayesian Estimation is also adopted as the alternative estimation of Lasso and Elastic Net Regression. The potential predictors include the GDP growth and the bilateral exchange rate of Thai trading partners. I use exports growth data from 2002:1 – 2019:4.

Results/Conclusion – For the frequentist approach, the elastic net regression with the weighting parameter set at 0.8 is the best performing model. The predicted series deviates from the actual series after the year 2017, which is attributable to the global trade shock resulting from US-China trade conflicts. Bayesian elastic net with 0.8 alpha weighing tops its frequentist equivalents in terms of MAPE and RMSE but returns a less satisfying forecast of the course of growth figures. Finally, regularization-based models largely outperform the AR (1) model in the forecast phases. Yet, regularization-based models are not substantially different from the factor model.

Keywords – Lasso, Adaptive Lasso, Elastic Net, and Adaptive Elastic Net, Bayesian Estimation

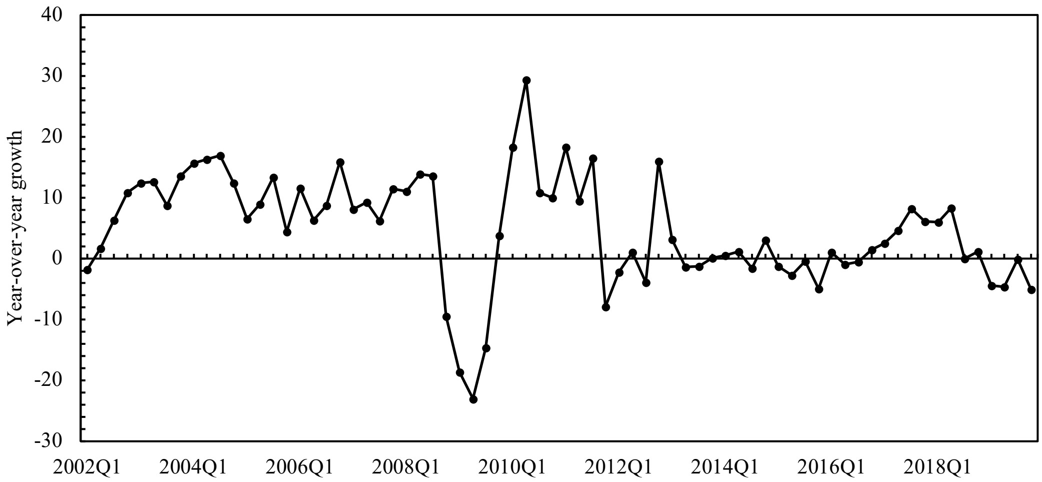
I. INTRODUCTION

A. Background

Exports have played a significant role in the Thai economy since the 1997 Asian financial crisis. (Sussangkarn and Nikomborirak, 2012) According to the National Account, the share of the real exports of goods to real GDP increased to 66.0 percent in 2008 from 45.3 percent in 1999. Even though export share to total GDP started to fall somewhat after the Global Financial Crisis, it remained high at 57.7 percent in 2018.

Exports projections have considerable economic implications. The accurate forecast allows a policymaker to project trade balance and overall current account balance, which have substantial implications to the economy, especially emerging economies that rely heavily on the external sector such as Thailand. A precise forecast of export growth allows policymakers to pre-emptively adjust macroeconomic stabilization policies, which will shelter the economy from an over-inflated economy or recession. Furthermore, the model helps detect some irregularities and misalignments in external sectors, allowing policymakers to react prudently and promptly to avoid adverse shocks such as severe capital outflows, exchange rate volatility, and foreign reserves mismanagement. On the other hand, the ability to project long-term exports growth trends would guide the long-term economic reforms. It provides a starting point for investigating the determinants of long-term export growth, helping policymakers identify and tackle the root causes of structural impediments.

Thai exports experienced booms and busts cycle throughout 2002:1 – 2019:4. During the early phases of the samples, export growth recovered from the previous downturn and stabilized at the elevated rate around 10 to 20 percent year-over-year. The buoyed growth had been observed across destinations and products. From categorical aspects, the buoyed headline was driven by the exports of computers and electronics, especially integrated circuited, and automobile and parts. From destination aspects, the export growth was driven by manufacturing-export-oriented countries which participated in the rapid expansion of China's manufacturing supply chain.



Source: Thailand Ministry of Commerce

Figure 3.1 Thai exports (quarterly data in year-over-year percentage growth)

During the year 2003-2007, export growth had recovered and stabilized at an elevated rate around 10 to 20 percent year-over-year. The buoyed growth had been observed across destinations and products, especially exports of agro-industry and manufacturing sectors. From categorical aspects, the steady growth was driven by the exports of computers and electronics, especially integrated circuited, and automobile and parts. From destination aspects, the export growth was driven by manufacturing-export-oriented countries which participated in the rapid expansion of China's manufacturing supply chain.

The next milestone is the breakthrough of the Global Financial Crisis (GFC) during 2008-2009. The global economy experienced an unprecedently sharp recession and the unparalleled financial turmoil in the global financial markets and banking systems. The GFC was catalysed by the severe slump in the US housing market and rolled out through the complex interconnection between the economy and financial markets. As global demand retreated, Thai export growth dropped sharply across every destination and category.

Export growth recovered and peaked out in 2011. Afterward, exports slowed down and stabilized at the lower level since the wake of the GFC. This headline reflected a subdued and vulnerable recovery in global demand as major countries experienced a wide range of hindrance: The eurozone suffered from the debt crisis since 2012, the deadly earthquake hit the north-eastern part of Japan in 2011, the wave of protectionism and deglobalization since 2018 was triggered by the reverse international policies of the US. Furthermore, Thai exports have suffered from structural problems. Many exporters cannot adapt to the fast-changing trends in the global economy. Some failed to move to a superior position on the global supply chain. These structural problems implied the permanent deceleration of export growth looking forward. (Apaitan, Ananchotikul, and Disyatat, 2017).

B. Research Objectives

The purpose of this paper is to investigate the accuracy of the model of Thai exports during 2002:1-2019:4 based on several regularization methods and to conduct an accuracy comparison. My contribution is to equip the traditional econometric forecasting strategies with the regularization techniques, which should help preserve parsimony and improve the forecasting accuracy of the model. Also, regularization techniques would theoretically mitigate the bias-variance trade-off compared to standard least square models. Regularization would help improve the efficiency of the estimates and forecasts, without adding to their biasedness to some extent.

This paper's objective is divided into three parts: Firstly, the first objective is to review forecasting strategies and regularization techniques to establish the candidate regularization techniques. Secondly, the intermediate and the advanced objectives focus on quantitively performing selected regularization techniques on the data set of Thai exports and the rich set of predictors. For the intermediate goal, I conduct the Ordinary Least Square Estimation of Lasso Regression, Adaptive Lasso Regression, Elastic Net Regression, and Adaptive Elastic Net Regression. For the advanced goal, I conduct the same regularization techniques using the Bayesian Estimation Method (BAY). Several accuracy measures would be adopted to compare performances.

C. Contributions

This paper will complement Korobilis (2013) and Stankiewicz (2015), who applied Bayesian techniques on elastic net estimation and forecasting as I research Thai export data. Korobilis (2013) explored the performance of lasso, elastic net, and adaptive elastic net with t-prior and Jeffreys before predicting macroeconomic variables, comparing with factor models. The author applied the techniques mentioned above on the U.S. macroeconomic data from 1959-2010 and contrasts different Bayesian shrinkage methods of factor models of forecasting. The study reveals that specific Bayesian-estimated elastic net models are superior to factor models.

Stankiewicz (2015) adopted Bayesian techniques to estimates lasso, adaptive lasso, elastic net, and adaptive elastic net based on the macroeconomic data from the Euro area from the first quarter of 1970 to the last quarter of 2013. The study revealed that the elastic net yielded the best performance, following by adaptive elastic net, lasso, and adaptive lasso. The result indicated the evidence against adaptive shrinkage. The study also asserts that the elastic net becomes even more superior to other methods when the number of predictors considerably exceeds that of the observation.

The remainder of this dissertation is organized as follows. Section 2 presents the theoretical setup of regularization techniques applied in this paper and reviews existed literature regarding exports forecast. While, section 3 discusses methodologies and the dataset and section 4 reports the results of the forecasting exercises. Lastly, Section 5 concludes this paper and directions for future work.

II. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

In this dissertation, I adopt several regularization techniques to estimate the linear regression model in the following generic form;

$$y = \beta_0 + \sum_{j=1}^{p} \beta_j x_j + \varepsilon \tag{2.1}$$

where y denotes the dependent variable, x_j denotes the predictors where $j = 1,2,\dots,p$, and $\varepsilon \sim N(0, \sigma^2)$ represents random disturbance or error term of the model. In this section, I present the mathematical specification and intuition of each regularization technique.

A. Regularization techniques

Ridge Regression

Ridge regression belongs to the class of regularization methods which aims at achieving the more stable estimates and avoiding overfitting problems. It "artificially" reduces the magnitude of the estimates of coefficients by imposing a penalty on each coefficient. Mathematically, the ridge regression minimizes the sum of squares of error plus penalty terms where coefficients are constrained to zero. The ridge regression coefficient, $\widehat{\beta^R}$ are estimated by minimizing the following sum of squares:

$$\sum_{i=1}^{n} (y_i - \beta_0^R - \sum_{j=1}^{p} \beta_j^R x_{ij})^2 + \lambda \sum_{j=1}^{p} (\beta_j^R)^2 \tag{2.2}$$

where $\lambda \geq 0$ is the tuning parameter that determines the strength of the regularization penalty.

For the value $\lambda = 0$, there is no regularization effect, and the estimates of the ridge regression coefficients are identical to the estimates of the coefficients obtained using the OLS method. As the value of λ increases, the weight put on the regularization penalty enlarges and the coefficient estimates approach zero. Unlike the OLS method, which generates only one set of coefficient estimates, the ridge regression method yields a set of coefficient estimates for each value of λ (James et al., 2013). Therefore, determining the fair value of λ is extremely important. The parameter λ is determined by the method of cross-validation.

Ridge regression is considered superior to the OLS regression in the aspect of balancing trade-offs between assessment bias and its variance. Ridge regression encompasses OLS regression as the special case where $\lambda = 0$. In this case, the coefficient estimates are unbiased, but the variance of the coefficient estimates are high. As the parameter λ increases, the flexibility of the regression model shrinks, which leads to a decrease in variance at the expense of higher bias. One can achieve optimal bias-variance trade-off by choosing the value of λ that minimizes mean squared error (MSE), which is a function of the variance and square of the bias.

Lasso Regression

The main disadvantage of ridge regression is the reflection in the impossibility to drop any predictors. As the value of parameter λ increases, the coefficients will move towards zero but will not be equal to zero[1]. Lasso regression tackles this limitation by changing the functional form of the penalty terms from sum of squares to sum of the absolute value of coefficient estimates. The lasso regression coefficients $\widehat{\beta^L}$ are estimated by minimizing the following terms:

$$\sum_{i=1}^{n}(y_i - \beta_0^L - \sum_{j=1}^{p} \beta_j^L x_{ij})^2 + \lambda \sum_{j=1}^{p} |\beta_j^L|. \tag{2.3}$$

As the value of tuning parameter λ increases, the lasso regression will drop some predictors since the estimates of their coefficients become zero. Thus, the lasso regression could contain any number of predictors depending on the value of the tuning parameter λ. In contrast, the estimates of the ridge regression coefficients would be nonzero. As a result, the ridge regression model will always contain all predictors. Noted that when $\lambda = 0$, the estimates of the lasso regression coefficients are identical to that of the OLS method. On the other hand, for a sufficiently large value of the parameter λ, estimate lasso regression contains no predictor as the estimates of the coefficients are equal to zero.

Adaptive Lasso Regression

Under some conditions, the estimates of Lasso regression coefficient lack consistency. This happens when the estimation procedure lacks the so-called oracle property. With the oracle properties, the oracle procedure could identify the true support and hold the optimal estimation rate. Oracle property ensures the consistency of the estimates and allows more flexibility in the degree of shrinkage. The adaptive lasso regression has the oracle property in a classical asymptotic situation, where the number of parameters $p \in N$ is fixed, and the number of observations goes to infinity $n \rightarrow \infty$.

Fan and Li (2001) believes that lasso does not have the oracle property. The problem with the lasso method is as follows. If the tuning parameter λ is large, the lasso draws non-zero coefficients too strongly towards zero. Conversely, if λ is small, lasso estimates too many coefficients as non-zero and is not consistent in estimating support. Meinshausen and Bühlmann (2006) showed that the latter case occurs when λ is chosen concerning the optimal predictive capabilities of the model (e.g. by cross-validation, Friedman et al., 2001).

[1] The only exception is the case when $\lambda \rightarrow \infty$

Meinshausen and Bühlmann (2006) have also shown that there are indeed examples of a regression matrix X in which lasso cannot have the oracle property for any choice of λ.

Based on the discussion in the previous paragraph, a natural solution is offered: an adaptive $\ell 1$-penalty, which would push small coefficients to zero with a higher penalty, and on the contrary, would penalize significant coefficients only slightly. Zou (2006) proposed an adaptive lasso in the following way:

$$\sum_{i=1}^{n}(y_i - \beta_0^{AL} - \sum_{j=1}^{p}\beta_j^{AL}\, x_{ij})^2 + \lambda\sum_{j=1}^{p} w_j\left|\beta_j^{AL}\right| \tag{2.4}$$

where $w_j \geq 0$ for each, $j = 1,.2,..,p$ denotes the heterogeneous adaptive weights across predictors. Conventionally, w_j are typically estimated from the data. I adopt the most common choice of weights as the inverse of ridge regression coefficient estimates, that is, $w_j = 1/\hat{\beta}^R$. Zou (2006) showed that under mild conditions of regularity, the adaptive lasso has the property of an oracle.

Elastic Net

Zou and Hastie (2005) proposed Elastic net which is a combination of lasso regression and ridge regression. Elastic net regression penalizes both the $\ell 1$- and $\ell 2$-norm of coefficients as follows,

$$\sum_{i=1}^{n}(y_i - \beta_0^{EN} - \sum_{j=1}^{p}\beta_j^{EN}\, x_{ij})^2 + \lambda\left(\alpha\sum_{j=1}^{p}|\beta_j^{EN}| + \frac{(1-\alpha)}{2}\sum_{j=1}^{p}\left(\beta_j^{EN}\right)^2\right) \tag{2.5}$$

whereas α indicates the relative weight of tuning parameter in favor of lasso penalty. Zou and Hastie (2005) mentioned many of the advantages that the elastic net has over the lasso method. The elastic net can be seen as a stabilized version of the lasso method in the same sense as the ridge regression by the stabilized version of the classical OLS estimate. The minimization problem of the objective function (2.5) is strictly convex due to $\ell 2$-penalty. The solution is therefore unambiguous, but again analytically inexpressible. Most of the algorithms used to find lasso estimates can be used with only minor modifications.

Adaptive elastic net

Adaptive elastic net is the adaptive variant of elastic net. Zou and Zhang (2009) designed an adaptive elastic net with a predominant $\ell 1$-norm and showed that it has oracle properties. The oracle property allows more flexibility in the degree of shrinkage. The multi-step adaptive elastic net can be defined by simply adding a $\ell 2$-penalty to the multi-step adaptive lasso method as follows.

$$\sum_{i=1}^{n}(y_i - \beta_0^{AEN} - \sum_{j=1}^{p}\beta_j^{AEN}\, x_{ij})^2 + \lambda\left(\alpha\sum_{j=1}^{p}w_j|\beta_j^{AEN}| + \frac{(1-\alpha)}{2}\sum_{j=1}^{p}\left(\beta_j^{AEN}\right)^2\right). \tag{2.6}$$

As I assign λ_1 the positive value, the estimation holds the advantage of lasso regression in terms of predictor shrinkages. Furthermore, with oracle properties and flexibility in terms of penalty factors, it tends to give better forecasting performances (Stankiewicz, 2015). However, the adaptive elastic net is not universally superior to lasso and ridge regression. Performance depends on the number of selected predictors. Namely, if the precise number of variables in the model is relatively small, relative penalty in favour of lasso regression may correctly select the relevant predictors. In contrast, if the optimal model contains a large number of variables, the inclination to ridge regression will give a better result since the coefficient estimate of none of predictors will be equal to zero. Since the exact number of predictors that the final model should contain is never known in advance, the method of cross-validation can determine which type of regression needs to be applied to a particular relationship.

B. Bayesian Estimation

Bayesian estimation of regularization techniques has two merits. Firstly, Bayesian estimation gives the posterior distribution of the parameters which provides reliable standard errors for statistical inference. Secondly, its framework allows us to put a prior on the penalty parameters. (Stankiewicz, 2015).

Instead of conducting the Gibb sampling, I adopt the alternative algorithm so-called the Empirical Bayesian Lasso and Empirical Bayesian Elastic Net Regression proposed by Cai et al. (2011) and Huang et al. (2015). This alternative algorithm does not require the full MCMC simulation, so it helps reduce the computational burden. Cai et al. (2011) and Huang et al. (2015) use hierarchical priors which preserve the distributional characteristics of Laplace distribution where a high probability mass is located at zero and both tails are heavy, which is consistent with the prior belief that most predictors could be dropped but some relevant predictors should have non-zero coefficients. These priors are conditional on variance parameter in the same fashion as that of Park and Casella (2008).

Firstly, Eq (2.1) is rewritten as $y = \mu + \sum_{j=1}^{p} \beta_j x_j + \varepsilon$. For Bayesian Lasso Regression, the prior probability distribution of beta coefficients follows independent normal distribution with mean zero and variance σ_j^2: $\beta_j \sim N(0, \sigma_j^2)$ where $j = 1, \dots, p$. On another layer, the variance σ_j^2 follow independent exponential distribution with hyperparameter λ which is the counterpart of the shrinkage parameter in Lasso Regression. The value of the hyperparameter could be selected by the cross-validation process. Based on this specification, the prior distributions of beta coefficients finally follow $p(\beta_j) = \sqrt{0.5\lambda}\exp(-\sqrt{2\lambda}|\beta_j|)$ which is compatible with Lasso penalty in Eq. (2.3).

For Bayesian Elastic Net Regression, Huang et al. (2015) proposes two hierarchical level priors. The first level is the prior probability distribution of beta coefficients which follows the independent normal distribution with mean zero and variance σ_j^2: $\beta_j \sim N(0, \sigma_j^2)$. The second level is the prior for the variance σ_j^2. Huang et al. (2014) decomposes the inverse of variance as $1/\sigma_j^2 = \lambda_1 + 1/\tilde{\sigma}_j^2$. The authors assume that, $\tilde{\sigma}_j^2$ follows a generalized Gamma distribution: $f(\tilde{\sigma}_j^2) = c(\lambda_1\tilde{\sigma}_j^2 + 1)^{-0.5}\exp(-\lambda_2\tilde{\sigma}_j^2)$. The merit of this specification is that, for any given positive value of λ_1, any given non-negative value of λ_2 and specific value of constant c, the prior distribution of β_j is proportional to $\exp(-0.5\lambda_1\beta_j^2 - \sqrt{2\lambda_2}|\beta_j|)$ which is compatible with the Elastic Net Penalty in Eq. (2.5).

C. Literature Review

The majority of economic literature adopts basic time-series methods to forecast economic variables. One of the very first research dates back to Mahmoud, Motwani and Rice (1990). In this seminal paper, the authors found that the simple time series techniques, such as the single exponential smoothing and the Carbone-Makridakis method, produce more satisfying results than the econometric model can predict.

Many following pieces of literature test the predictability of the more developed time series models, such as Autoregressive Models. For instance, Kumar and Gupta (2010) develop the Autoregressive Integrated Moving Average (ARIMA) model to forecast exports of industrial goods from Punjab, one of India's economically essential states. The authors adopt the Box-Jenkins approach to specify the structure of the ARIMA model and report mildly satisfying results. The more recent research augments the ARIMA model with the explanatory variables. These variables are selected based on economic theories and statistical observations. For example, Kongcharoen and Kruangpradit (2013) forecast Thai export value with the Autoregressive Integrated Moving Average with the Explanatory Variable (ARIMAX) Model.

Later on, economists and practical forecasters recognize the endogeneity between trade values and their determinants. It warrants the use of the Vector Autoregression Model (VAR) and its variants. VAR analysis allows the dynamic influence of across variables in the system. So, it tends to be superior to a univariate AR approach for capturing the long-run dynamics of the variables. Ugur (2010) forecasts the quarterly GDP growth of the Turkish economy during 1994:1 and 2005:4. The results suggest that the structure of the model considerably captures the data dynamics rather well and produces the economically rational prediction.

Recent literature explores the application of machine learning methodologies on economic forecasting. For instance, Co and Boosarawongse (2007) contrast the Artificial Neural Networks (ANNs) with various statistical techniques in forecasting Thailand rice exports. Theoretically, ANNs develop an internal representation of the relationship between the variables of interest. This representation is highly flexible as the ANN comprises the non-linear mapping system and does not assume any distributional nature of the data. Therefore, ANNs are expected to produce a more accurate forecast of economic data than traditional econometric models, which impose distributional assumptions on the data. The result shows that ANNs shows satisfying forecasting performance of the testing data as ANNs can capture the non-linearity and seasonality pattern of the data. Recent work of Urrutia, Abdul, and Atienza (2019) adopt the Bayesian ANNs to forecast Philippines imports and exports. Compare to its frequentist counterparts, the Bayesian ANNs allow the researcher's information inputs and produce the full distribution of results that allow robust inferences. The result suggests that the Bayesian ANNs significantly beat the traditional ARIMA model. The Paired T-test find no statistically significant difference between actual and forecast series of import and export value.

III. METHODOLOGY

The main dataset for Thai exports comprises quarterly Thai exports year-over-year growth from 2002:1 to 2019:4. The data is normalized by using the z-transformation. The data set is separated into a training set (Fitted phases) to estimate the regression model and a testing set (Forecasted phases) to produce a forecast. The training set comprises the data from 2002:1 to 2013:4 (48 data points). The testing set comprises the data from 2014:1 to 2019:4 (24 data points). The ratio between the training set and testing set becomes 2:1. To assess the consistency and stability of forecasting results, the data series is also partitioned further. For training set, I partitioned the series into two periods: before the Global Financial Crisis (GFC) (2002:1 – 2007:4) and since the GFC (2008:1 – 2013:4). For testing set, I also partitioned the series into two forecasting phases: 2014:1 – 2016:4 and 2017:1 – 2019:4).

I select two groups of candidate predictor. The first group represents the aggregate income of trading partners. I collect the data series of quarterly real GDP growth in a year-over-year basis. As the income effect could be persistent, one-quarter lag and two-quarters lag of real GDP growth are also added to potential predictors. This data set covers 44 trading partners across Asia, Europe, North and South America, and Africa. The second group comprises relative trading prices. This group contains the bilateral indirect exchange rate between Thai Baht and 35 trading partner currencies. The candidates are selected based on their relevance to Thai exports and the data availability. All predictors are also normalized using z-transformation.

Table 3.1. List of trading partner countries which are included in the "aggregate income" group of predictors

ARG	Argentina	FIN	Finland	LVA	Latvia
AUS	Australia	FRA	France	MEX	Mexico
AUT	Austria	GBR	Great Britain	NLD	Netherlands
BEL	Belgium	GRC	Greece	NOR	Norway
BGR	Bulgaria	HUN	Hungary	NZL	New Zealand
BRA	Brazil	IDN	Indonesia	POL	Poland
CAN	Canada	IND	India	PRT	Portugal
CHE	Switzerland	IRL	Ireland	ROU	Romania
CHL	Chile	ISL	Iceland	SVK	Slovakia
CHN	China	ISR	Israel	SVN	Slovenia
CZE	Czech Republic	ITA	Italy	SWE	Sweden
DEU	Germany	JPN	Japan	TUR	Turkey
DNK	Denmark	KOR	Republic of Korea	USA	United States of America
ESP	Spain	LTU	Lithuania	ZAF	South Africa
EST	Estonia	LUX	Luxembourg		

Table 3.2. List of currencies which are included in the "relative trading price" group of predictors

AED	UAE Dirham	IDR	Rupiah	PHP	Philippine Peso
AUD	Australian Dollar	INR	Indian Rupee	PKR	Pakistan Rupee
BDT	Taka	KES	Kenyan Shilling	RUB	Russian Ruble
BND	Brunei Dollar	KHR	Riel	SAR	Saudi Riyal
CAD	Canadian Dollar	KRW	Won	SEK	Swedish Krona
CHF	Swiss Franc	KWD	Kuwaiti Dinar	SGD	Singapore Dollar
CNY	Yuan Renminbi	LAK	Kip	TWD	New Taiwan Dollar
CZK	Czech Koruna	MMK	Kyat	USD	US Dollar
DKK	Danish Krone	MXN	Mexican Peso	VND	Dong
EUR	Euro	MYR	Malaysian Ringgit	YEN	Japanese Yen
GBP	Pound Sterling	NOK	Norwegian Krone	ZAR	Rand
HKD	Hong Kong Dollar	NZD	New Zealand Dollar		

This dissertation compares forecasting accuracy between the regression which based on the following regularization methods: Lasso Regression, Adaptive Lasso Regression, Elastic Net Regression, and Adaptive Elastic Net Regression. These regularization techniques are conducted on two different estimation strategies: Ordinary Least Square Method (OLS) and Bayesian Estimation Method (BAY). For each pair of the regularization technique and the estimation strategy, I firstly conduct the regularization to select relevant covariates. I adopted ten-folded cross-validation techniques to select the values of penalty parameters for each pair. To ensure the robustness of OLS estimation, the cross-validation procedure was repeated 100 times, and the covariates which are selected for at least 50 out of 100 times were adopted. After the predictors are selected, I re-fit the final model with the OLS estimation and measure its predictive accuracy.

I adopt three widely adopted accuracy measures. The first two accuracy measures are Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) which are the straightforward comparison between actual and forecasted series. Furthermore, I recognize that the practical forecasting involves predicting future direction of export growth, whether the figure accelerates or decelerates from the previous period. Thus, I construct the Boolean series of actual and forecasted series, which takes the value of 1 if the figure accelerates and 0 otherwise. I adopt the Area Under the ROC Curve (AUC) which indicates the model's capacity of distinguishing acceleration and deceleration.

IV. ESTIMATION RESULTS

A. OLS Estimation of Selected Regularization Techniques

Firstly, I compare the forecasting performance of lasso and elastic net regression.[2] Among lasso and elastic net regression, I deem that the best performing model in the forecasted phases is the elastic net regression with the weighting parameter set at 0.8. This model's MAPE and RMSE are among the lowest (1.47 and 1.22, respectively). Furthermore, its prediction of the acceleration or deceleration of export growth is the most accurate, returning the highest AUC at 0.57. This result is consistent across the subperiod of forecasted phases.

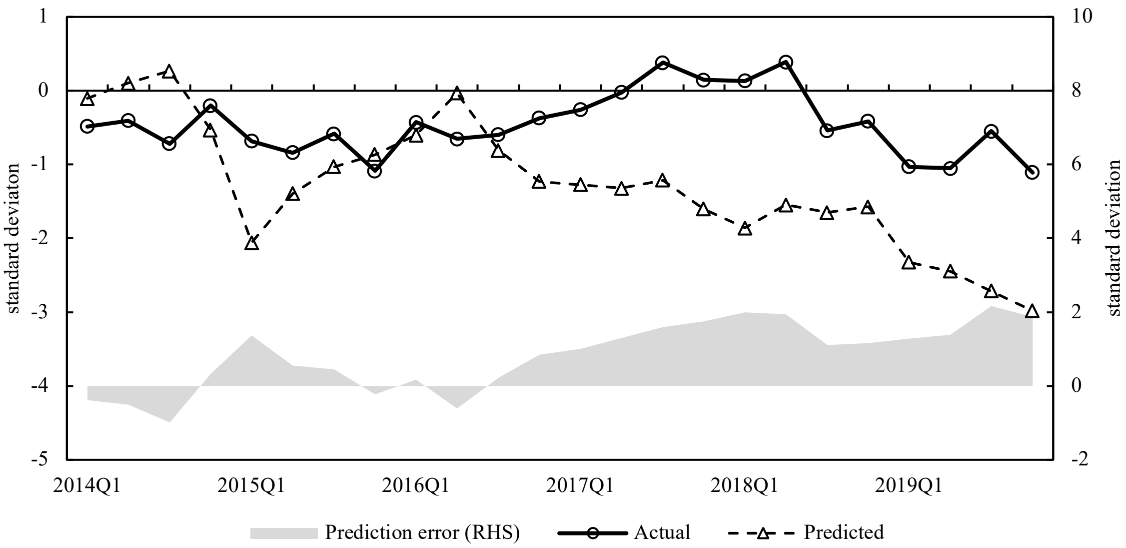
Figure 4.1 illustrates the forecasting accuracy of the best performing model across time. It shows that the predicted series is rather close to the original series during the year 2014-2016. However, the predicted series started to deviate from the actual series during the year 2017-2019. I conjecture that this result is attributable to the global trade shock resulting from US-China trade conflicts. It warrants more recent data points to retrain the model in order to capture this changing structure.

Next, I compare the forecasting performance of lasso and elastic net regression with the adaptive lasso and the adaptive elastic net regression. Theoretically, the adaptive lasso and adaptive elastic net hold oracle properties, which return the more consistent estimates compare to lasso and elastic net counterparts. However, empirical literature indicates that their adaptive counterparts do not strictly dominate lasso and elastic net models.

Table 4.1. Summary of accuracy measures for regression specifications based on selected regularization methods-Thai exports (quarterly data in year-over-year percentage growth)

			Elastic Net									
	Accuracy	L2	Alpha									L1
Data points	Measure		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
			Fitted phase									
Year 2002-2013	MAPE	0.00	0.00	0.00	0.24	0.44	0.29	0.30	0.23	0.29	0.28	0.28
Data points 1-48	RMSE	0.00	0.00	0.00	0.09	0.09	0.09	0.09	0.10	0.12	0.12	0.12
	AUC	1.00	1.00	1.00	0.89	0.94	0.96	0.96	0.92	0.89	0.89	0.89
2002-2007	MAPE	0.00	0.00	0.00	0.30	0.60	0.38	0.39	0.28	0.34	0.34	0.34
1-25	RMSE	0.00	0.00	0.00	0.09	0.07	0.08	0.08	0.09	0.11	0.11	0.11
	AUC	1.00	1.00	1.00	0.90	0.98	1.00	1.00	0.95	0.90	0.90	0.90
2008-2013	MAPE	0.00	0.00	0.00	0.26	0.49	0.32	0.33	0.24	0.29	0.29	0.29
26-48	RMSE	0.00	0.00	0.00	0.09	0.07	0.08	0.08	0.09	0.11	0.11	0.11
	AUC	1.00	1.00	1.00	0.89	0.96	0.98	0.98	0.94	0.90	0.90	0.90
			Forecasted phase									
2014-2019	MAPE	0.93	0.97	1.02	1.44	1.75	1.76	1.69	2.30	1.47	1.86	1.86
49-72	RMSE	10.11	13.53	21.15	0.77	2.16	2.41	2.46	1.68	1.22	1.22	1.22
	AUC	0.43	0.44	0.33	0.39	0.48	0.44	0.44	0.35	0.57	0.57	0.57
2014-2016	MAPE	0.91	0.96	0.98	2.05	2.59	2.61	2.46	3.72	2.08	2.85	2.85
49-60	RMSE	7.74	16.49	14.67	0.80	1.38	1.50	1.53	0.77	0.65	0.65	0.65
	AUC	0.47	0.13	0.34	0.07	0.37	0.37	0.37	0.34	0.34	0.34	0.34
2017-2019	MAPE	0.96	0.98	1.06	0.84	0.90	0.91	0.91	0.88	0.87	0.87	0.87
61-72	RMSE	12.03	9.70	26.07	0.73	2.72	3.05	3.12	2.25	1.60	1.59	1.59
	AUC	0.34	0.63	0.27	0.63	0.65	0.54	0.54	0.34	0.88	0.88	0.88

[2] The list of selected variables and configuration of penalty parameter are shown in Appendix



Source: Bank of Thailand, Bank of International Settlement

Figure 4.1. Actual and predicted series of Thai quarterly year-over-year export growth from the Elastic Net Regression (Alpha = 0.8)

Table 4.2 summarizes the forecasting performance of the adaptive lasso and adaptive elastic net regression. Among the adaptive models, adaptive lasso exhibits the lowest MAPE, while adaptive elastic net with the weighting parameter set at 0.2 returns the lowest RMSE. In contrast, the adaptive elastic net with the weighting parameter set at 0.3-0.6 is the best performing model for forecasting the acceleration or deceleration of export growth. However, these adaptive models' performance does not considerably outperform the best performing model without adaptive shrinkage. For example, MAPE and RMSE of adaptive elastic net given the weighting parameter of 0.2 might be slightly lower than the best performing model without adaptive shrinkage; the prediction of growth acceleration is less accurate. Given an additional computational burden, I choose the model without adaptive shrinkage to represent OLS-estimated models in the following comparison with Bayesian models and econometric models.

B. Bayesian Estimation of Selected Regularization Techniques

Table 4.3 exhibits a summary of the accuracy of the Bayesian Lasso Regression and Bayesian Elastic Net Regression. Among the Bayesian models, Bayesian Elastic Net Regression with the hyperparameter alpha 0.8 outperform others during the forecasted phases. From table 4.3, the Bayesian Elastic Net Regression with alpha 0.8 is the most accurate forecasting model based on MAPE and RMSE, which is consistent across subperiods. Meanwhile, it performs relatively fair in predicting the acceleration or the deceleration of exports growth more accurately, with the AUC at 39%.

Figure 4.3 compares the predicted export growth from Bayesian Elastic Net with alpha at 0.8 with the actual export growth during forecasting phase. In contrast to frequentist Elastic Net, Bayesian Elastic Net gives a more accurate forecast during the recent period from year 2017-2019. This result is expected given that the pair suggests the substantially different sets of selected predictors.

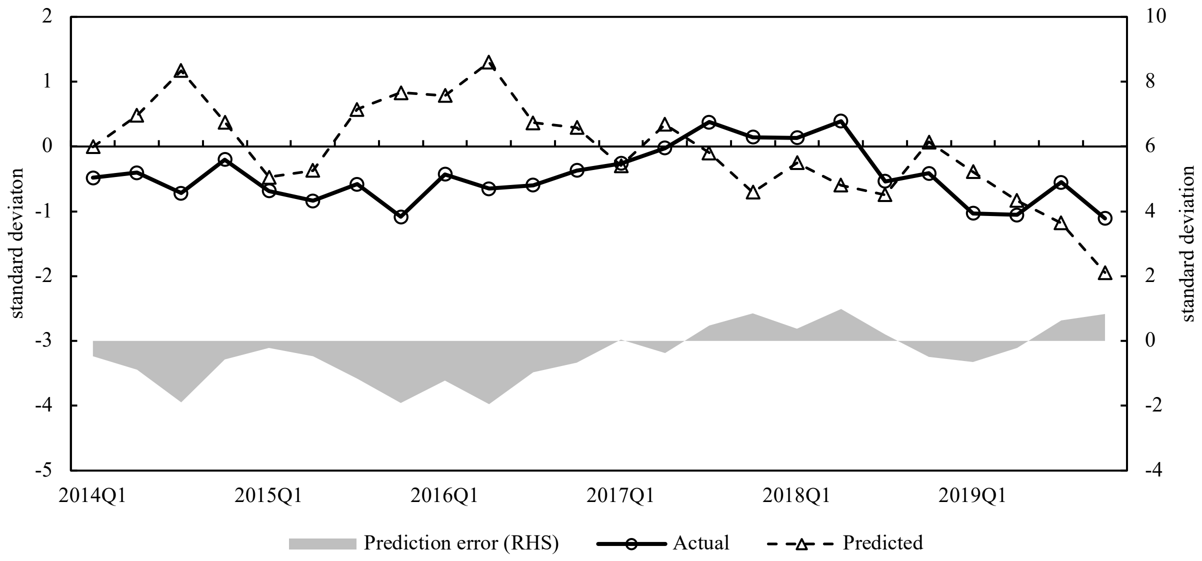
Table 4.4 compares the predictability of the best performing OLS-based shrinkage model and the best performing Bayesian shrinkage model. The result shows that Bayesian Elastic Net with the weighting parameter of 0.8 beats its frequentist counterparts based on MAPE and RMSE in the forecasted period. However, the OLS-estimated Elastic Net is better for predicting future growth figures' direction as indicated better AUC. I note that the Bayesian shrinkage returns the slightly more parsimonious model as it shrinks 167 potential predictors down to 30 predictors, compared to its frequentist counterparts, which shrinks the number of predictors down to 35 predictors.

Table 4.2. Summary of accuracy measures for regression specifications based on selected adaptive regularization method-Thai exports (quarterly data in year-over-year percentage growth)

Data points	Accuracy Measure	Adapt. L2	Adaptive Elastic Net Alpha 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Adapt L1
					Fitted phase							
Year 2002-2013	MAPE	0.00	0.15	0.12	0.28	0.28	0.28	0.30	1.30	1.30	1.30	0.85
Data points 1-48	RMSE	0.00	0.05	0.07	0.10	0.10	0.10	0.10	0.11	0.11	0.11	0.11
	AUC	1.00	0.96	0.95	0.91	0.91	0.91	0.91	0.93	0.93	0.93	0.93
2002-2007	MAPE	0.00	0.19	0.16	0.25	0.25	0.25	0.25	0.32	0.32	0.32	0.76
1-25	RMSE	0.00	0.05	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.10
	AUC	1.00	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
2008-2013	MAPE	0.00	0.10	0.08	0.31	0.31	0.31	0.35	2.29	2.29	2.29	0.95
26-48	RMSE	0.00	0.05	0.07	0.11	0.11	0.11	0.11	0.12	0.12	0.12	0.12
	AUC	1.00	0.96	0.96	0.87	0.87	0.87	0.87	0.91	0.91	0.91	0.91
					Forecasted phase							
2014-2019	MAPE	0.93	2.67	1.45	1.76	1.76	1.76	1.48	2.04	2.04	2.04	1.28
49-72	RMSE	10.11	2.75	1.57	2.06	2.06	2.06	2.05	1.98	1.98	1.98	1.93
	AUC	0.43	0.30	0.48	0.57	0.57	0.57	0.57	0.48	0.48	0.48	0.53
2014-2016	MAPE	0.91	4.38	1.72	2.61	2.61	2.61	2.05	3.17	3.17	3.17	1.66
49-60	RMSE	7.74	3.68	1.91	1.32	1.32	1.32	1.29	1.35	1.35	1.35	1.35
	AUC	0.47	0.19	0.45	0.45	0.45	0.45	0.45	0.37	0.37	0.37	0.37
2017-2019	MAPE	0.96	0.96	1.17	0.91	0.91	0.91	0.91	0.90	0.90	0.90	0.90
61-72	RMSE	12.03	1.26	1.13	2.60	2.60	2.60	2.60	2.46	2.46	2.46	2.37
	AUC	0.34	0.34	0.54	0.73	0.73	0.73	0.73	0.65	0.65	0.65	0.73

Table 4.3. Summary of accuracy measures for regression specifications based on selected Bayesian regularization method-Thai exports (quarterly data in year-over-year percentage growth)

		Bayesian Elastic Net Alpha 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Bayesian L1
				Fitted phase							
Year 2002-2013	MAPE	0.81	1.82	0.75	0.53	0.66	0.50	0.50	0.92	1.32	0.55
Data points 1-48	RMSE	0.31	0.52	0.28	0.25	0.25	0.24	0.22	0.40	0.52	0.19
	AUC	0.85	0.68	0.89	0.91	0.87	0.85	0.85	0.75	0.68	0.91
2002-2007	MAPE	0.90	0.85	0.84	0.66	0.53	0.58	0.54	0.68	1.06	0.90
1-25	RMSE	0.21	0.34	0.24	0.22	0.23	0.20	0.18	0.26	0.35	0.17
	AUC	0.85	0.79	0.85	0.89	0.89	0.85	0.85	0.85	0.79	0.94
2008-2013	MAPE	0.73	2.79	0.67	0.40	0.80	0.42	0.46	1.17	1.58	0.21
26-48	RMSE	0.38	0.64	0.32	0.27	0.27	0.27	0.25	0.50	0.64	0.20
	AUC	0.87	0.58	0.91	0.91	0.87	0.87	0.87	0.68	0.58	0.91
				Forecasted phase							
2014-2019	MAPE	3.10	2.03	2.92	57.75	3.47	12.78	3.70	1.20	1.69	1.97
49-72	RMSE	0.57	0.44	0.80	0.68	0.67	0.67	0.59	0.43	0.42	1.10
	AUC	0.29	0.44	0.35	0.48	0.48	0.44	0.48	0.39	0.43	0.23
2014-2016	MAPE	3.53	1.97	3.30	113.06	4.79	23.68	3.41	0.83	1.42	2.31
49-60	RMSE	0.71	0.41	1.00	0.85	0.85	0.84	0.73	0.41	0.36	1.31
	AUC	0.27	0.27	0.27	0.27	0.18	0.18	0.27	0.18	0.34	0.07
2017-2019	MAPE	2.67	2.10	2.54	2.44	2.15	1.88	3.99	1.58	1.96	1.63
61-72	RMSE	0.38	0.46	0.52	0.45	0.42	0.43	0.39	0.45	0.48	0.83
	AUC	0.22	0.54	0.34	0.63	0.73	0.63	0.63	0.54	0.45	0.22



Source: Bank of Thailand, Bank of International Settlement

Figure 4.3. Actual and predicted series of Thai quarterly year-over-year export growth from Bayesian Elastic Net (Alpha = 0.8)

C. Comparing Selected Regularization-Based Model with Econometric and Factor-Based Model

In this section, I compare the forecasting performance of the frequentist and Bayesian regularized regression models with a simple ARMA model and a factor model with 4 principal components. To specify the ARMA structure, I conduct the grid search for the structure that yields the lowest AIC test statistics. The grid search suggests AR (1) structure. For the factor model, I conduct a principal component analysis to extract principal components from the datasets of potential predictors. I adopt the first four components which cumulatively account for 73.29% of the total variance of all potential predictors.

Table 4.4 suggests that the regularization-based models broadly outperform the AR (1) model in the forecasted phases. This result is consistent across the accuracy measures. On the other hand, regularization-based models are not considerably different from the factor model. On the basis of MAPE, the factor model outperforms both regularized models. On the other hand, the Bayesian-estimated Elastic Net Regression is more accurate on the basis of RMSE. Furthermore, the OLS-estimated Elastic Net Regression gives a more satisfying prediction of the direction of exports growth.

Table 4.4. Summary of accuracy measures for regression specifications based on selected Bayesian regularization method-Thai exports (quarterly data in year-over-year percentage growth)

Data points	Accuracy Measure	Elastic Net (alpha = 0.8) with OLS Estimation	Elastic Net (alpha = 0.8) with Bayesian Estimation	AR (1)	4 Principal Components
			Fitted phases		
Year 2002-2013	MAPE	0.29	0.92	6.90	2.03
Data points 1-48	RMSE	0.12	0.40	0.88	0.61
	AUC	0.89	0.75	0.43	0.68
2002-2007	MAPE	0.34	0.68	1.33	0.62
1-25	RMSE	0.11	0.26	0.45	0.45
	AUC	0.90	0.85	0.38	0.67
2008-2013	MAPE	0.29	1.17	12.48	3.44
26-48	RMSE	0.11	0.50	1.15	0.74
	AUC	0.90	0.68	0.44	0.64
			Forecasted phases		
2014-2019	MAPE	1.47	1.20	3.33	0.67
49-72	RMSE	1.22	0.43	0.46	1.02
	AUC	0.57	0.39	0.34	0.49
2014-2016	MAPE	2.08	0.83	4.87	0.49
49-60	RMSE	0.65	0.41	0.44	0.58
	AUC	0.34	0.18	0.27	0.35
2017-2019	MAPE	0.87	1.58	1.80	0.85
61-72	RMSE	1.60	0.45	0.49	1.32
	AUC	0.88	0.54	0.37	0.80

Source: Bank of Thailand, Bank of International Settlement

V. CONCLUSION

This study intends to leverage various regularization strategies to determine the specification of Thai export growth. I test the forecasting performance of multiple models developed on various regularization strategies and contrast these models with the Autoregressive model and factor model.

For the frequentist approach, the elastic net regression with the weighting parameter set at 0.8 is the best performing model during forecasted phases. The predicted series is rather close to the original series during the year 2014-2016. The predicted series deviates from the actual series after the year 2017, which is attributable to the global trade shock resulting from US-China trade conflicts. It warrants more recent data points to retrain the model to capture this changing structure of international trade. I also find that OLS-estimated adaptive shrinkage does not significantly outperform the non-adaptive counterparts.

Comparing OLS-based shrinkage models with Bayesian shrinkage models, the result shows that the Bayesian Elastic Net with the weighting parameter of 0.8 beats its frequentist counterparts based on MAPE and RMSE. However, the OLS-estimated Elastic Net is better for predicting future growth figures' direction as indicated better AUC. I note that the Bayesian shrinkage gives the more parsimonious model. Lastly, the regularization-based models broadly outperform the AR (1) model in the forecasted phases. Regularization-based models are not considerably different from the factor models, with the first four principal components of all potential predictors.

Forecasting exercise suggests that the regularization techniques return satisfying results. It could give a comparable performance with the factor models and even top the widely-used autoregressive models. It suggests that the regularization techniques could help improve the performance of practical forecasting and warrant further academic investigation. It is also possible to apply the regularization techniques with other types of forecasting models such as the VAR and Error Correction Model, which might enhance forecasting accuracy in some specific contexts.

REFERENCES

- Apaitan T., Ananchotikul N., and Disyatat P. (2017). Structural Transformation in Thailand: Perspective Through Product Innovation. Thesis at: *Puey Ungphakorn Institute for Economic Research.*, (72). Available at: https://www.pier.or.th/ (Accessed on 4 August 2020)
- Cai, X., Huang, A., and Xu, S. (2011). Fast empirical Bayesian LASSO for multiple quantitative trait locus mapping. BMC Bioinformatics 12, 211. Available at: https://www.researchgate.net/
- (Accessed on 4 August 2020)
- Co, H., and Boosarawongse, R. (2007). Forecasting Thailand's rice export: Statistical techniques vs. Artificial neural networks. Computers & Industrial Engineering 53: 610-627. Available at: https://www.sciencedirect.com (Accessed on 30 June 2020)
- Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96 (456),1348–1360. Available at: https://orfe.princeton.edu/ (Accessed on 30 June 2020)
- Friedman, J., Hastie, T. and Tibshirani, R. (2001). The elements of statistical learning, Berlin: Springer series in statistics Springer, pp. 43-92
- Huang, A., Xu, S., and Cai, X. (2015). Empirical Bayesian elastic net for multiple quantitative trait locus mapping. Heredity 114(1): 107-115. Available at: https://www.nature.com/articles/hdy201479
- (Accessed on 15 August 2020)
- James, G., Witten, D., Hastie, T., and Tibshirani R. (2013). Linear Model Selection and Regularization. In: An Introduction to Statistical Learning. Springer Texts in Statistics, 103. S New York: Springer. Available at: https://doi.org/10.1007/978-1-4614-7138-7_6 (Accessed on 8 August 2020)
- Kongcharoen, C. and Kruangpradit, T. (2013). Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX) Model for Thailand Export. Available at: https://forecasters.org (Accessed on 8 August 2020)
- Korobilis D (2013), Hierarchical shrinkage priors for dynamic regressions with many predictors, *International Journal of Forecasting*. Vol. 29(1), pp. 43-59.
- Kumar, G., and Gupta, S. (2010). Forecasting Exports of Industrial Goods from Punjab - An Application of Univariate Arima Model. *Annals of the University of Petrosani* 10(4):169-180. Available at: https://www.semanticscholar.org (Accessed on 12 August 2020)
- Mahmoud, E., Motwani, J. and Rice, G. (1990). Forecasting US exports: An illustration using time series and econometric models. Omega 18(4):375-382. Available at: https://www.sciencedirect.com/ (Accessed on 12 August 2020)
- Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *The annals of statistics,* 1436–1462. Available at: https://arxiv.org/pdf/math/0608017.pdf (Accessed on 15 July 2020)
- Park, T. and G. Casella (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, 103, 681–686. Available at: https://people.eecs.berkeley.edu (Accessed on 15 August 2020)
- Stankiewicz, S. (2015), Forecasting Euro Area Macroeconomic Variables with Bayesian Adaptive Elastic Net. Thesis at: *Department of Economics, University of Konstanz.*, (2015-12) Available at: https://kops.uni-konstanz.de (Accessed on 20 April 2020)
- Sussangkarn, C., and Nikomborirak, D. (2012), Thailand: Post-Crisis Rebalancing, Asian Economic Papers. *MIT Press - Journals*. Vol. 11(1), pp. 1-26

- Ugur, A. (2010). Import and Economic Growth in Turkey: Evidence from Multivariate VAR Analysis. Available at: https://www.u-picardie.fr/eastwest/fichiers/art68.pdf (Accessed on 13 August 2020)
- Urrutia, J. D., Abdul, A. M., and Atienza J. B. E. (2019) Forecasting Philippines imports and exports using Bayesian artificial neural network and autoregressive integrated moving average. *AIP Conference Proceedings* 2192:1. Available at: https://aip.scitation.org/doi/pdf/10.1063/1.5139185 (Accessed on 13 August 2020)
- Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B*, Royal Statistical Society 67(2): 301-320. Available at: https://web.stanford.edu/~hastie (Accessed on 20 July 2020)
- Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476): 1418–1429. Available at: http://users.stat.umn.edu (Accessed on 18 May 2020)
- Zou, H. and Hastie, T. (2012). Elastic-Net for Sparse Estimation and Sparse PCA. Available at: www.stat.umn.edu/~hzou. (Accessed on 22 May 2020)
- Zou, H. and Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of statistics*, 37(4), 1733. Available at: http://users.stat.umn.edu/~zouxx019/Papers/aenet.pdf (Accessed on 27 July 2020)

APPENDIX

Table A.1. Selected variables based on Elastic Net and Adaptive Elastic Net

	Elastic Net											Adaptive Elastic Net										
	L2										L1	L2										L1
alpha	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
G.ARG	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.AUS	*	*										*										
G.AUT	*	*										*										
G.BRA	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.DNK	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.EST	*	*	*									*										
G.FIN	*	*	*	*	*	*	*	*				*										
G.GRC	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.IDN	*	*	*	*	*	*	*	*	*			*	*									
G.IND	*	*	*									*										
G.ISL	*	*										*										
G.ISR	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*				
G.KOR	*	*	*	*	*	*	*					*	*	*								
G.LTU	*	*			*			*	*	*	*	*										
G.LVA	*	*										*										
G.MEX	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.NLD	*	*										*										
G.NOR	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.NZL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
G.POL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.ROU	*	*	*	*								*	*	*								
G.SVK	*	*	*	*	*	*	*	*	*	*	*	*										
G.SWE	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
G.TUR	*	*										*										
LG.BGR	*	*										*										
LG.BRA	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LG.CAN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LG.CHE	*	*										*										
LG.DNK	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LG.GBR	*	*	*	*	*							*	*	*								
LG.GRC	*	*										*										
LG.IND	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LG.IRL	*	*	*		*	*	*					*	*	*	*	*	*	*	*	*	*	*
LG.ISL	*	*	*		*	*	*	*				*	*	*	*	*	*	*	*	*	*	*
LG.ISR	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*				
LG.JPN	*	*	*	*	*	*	*	*				*	*	*	*	*	*	*	*	*	*	*
LG.LTU	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LG.LUX	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LG.NOR	*	*	*	*	*	*	*	*	*	*	*	*	*	*								
LG.NZL	*	*										*										
LG.ROU	*	*										*	*									
LG.SVK	*	*										*										
LG.SWE	*	*										*										
LG.USA	*	*	*									*	*									
LLG.ARG	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

LLG.AUS	*	*										*										
LLG.BGR	*	*	*									*	*	*								
LLG.CHE	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.CHL	*	*										*										
LLG.DNK	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.IDN	*	*										*										
LLG.ISL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.ITA	*	*										*										
LLG.JPN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.KOR	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.LTU	*	*										*										
LLG.LUX	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.LVA	*	*	*									*	*									
LLG.NLD	*	*										*										
LLG.NOR	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LLG.NZL	*	*	*	*	*							*										
LLG.POL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C.YEN	*											*		*								
C.MYR	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C.SGD	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*				
C.BND	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C.PHP	*	*	*	*	*	*						*	*	*	*	*	*					
C.INR	*	*										*										
C.CHF	*	*										*										
C.AUD	*	*	*	*								*										
C.SEK	*	*										*										
C.NOK	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C.CNY	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Number of variables	167	72	49	42	43	40	39	38	35	34	34	167	44	41	35	35	35	34	31	31	31	30

Note: Prefix “G.” denotes the GDP growth of the country whose abbreviated name follows. Prefix “L” denotes one-period lag while “LL” denotes two-period lag. Prefix “C.” denotes the foreign exchange rate of the currency whose abbreviated name follows.

Table A.2. Selected variables based on Bayesian Elastic Net

alpha	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	L1 1.0
G.ARG				*				*		*
G.BGR							*			
G.BRA	*	*	*	*	*	*	*			*
G.CHL		*		*		*				
G.DEU									*	
G.DNK	*	*	*	*	*	*	*	*		
G.FIN		*	*	*		*				
G.FRA									*	
G.GBR									*	
G.GRC				*						
G.HUN										*
G.IDN										*
G.IND								*		
G.ISR										*
G.ITA	*	*		*	*	*	*			
G.KOR			*							
G.LUX		*	*	*		*		*		
G.LVA					*					
G.MEX								*	*	*
G.NOR	*						*			
G.NZL					*			*		
G.POL	*	*	*	*				*		
G.SVK	*								*	*
G.SWE	*	*	*	*	*	*	*	*	*	*
G.TUR			*							
G.USA									*	
LG.BEL	*	*	*	*	*	*	*	*		*
LG.BRA					*	*			*	
LG.CAN	*	*	*	*	*	*	*	*	*	*
LG.CHE			*							
LG.CHL								*		
LG.DEU				*		*		*		
LG.GRC										*
LG.HUN			*	*				*	*	*
LG.IND	*	*	*	*	*	*	*	*	*	*
LG.IRL				*				*		
LG.ISR	*								*	*
LG.JPN			*					*	*	*
LG.LTU	*	*	*	*	*	*	*	*		
LG.LUX	*	*	*	*	*	*	*	*	*	*
LG.NOR	*				*		*			
LG.USA	*	*	*	*	*	*	*	*		
LLG.ARG	*	*	*	*	*	*	*	*	*	*
LLG.AUT								*	*	
LLG.CHE									*	
LLG.DNK	*	*	*	*	*	*	*	*	*	*
LLG.EST					*		*			
LLG.IDN								*	*	
LLG.ISL			*							
LLG.ITA			*							
LLG.JPN	*	*	*	*	*	*	*	*	*	*
LLG.KOR		*							*	*
LLG.LUX								*	*	*
LLG.LVA						*	*	*		
LLG.NOR	*	*	*	*	*	*	*	*	*	*
LLG.NZL	*								*	*
LLG.POL	*	*	*	*	*	*	*	*		
LLG.TUR									*	*
LLG.USA	*			*				*		
C.SGD	*			*						
C.BND	*	*	*	*	*	*	*	*	*	*
C.AUD								*		
C.CNY										*
Number of variables	23	20	24	26	20	21	20	30	25	25

Table A.3. Configuration of hyperparameters (penalty parameter, lambda)

	L2										L1
alpha	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Search range						[0.01,1000]					
Elastic Net											
Median	0.0501	0.0501	0.0316	0.0251	0.0158	0.0126	0.0100	0.0100	0.0100	0.0100	0.0100
Min	0.0100	0.0158	0.0158	0.0126	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
Max	0.2512	0.1000	0.0631	0.0398	0.0316	0.0251	0.0200	0.0200	0.0158	0.0158	0.0631
Adaptive Elastic Net											
Median	0.5012	0.3162	0.1585	0.1259	0.1000	0.0794	0.0631	0.0501	0.0501	0.0398	0.0398
Min	0.1995	0.1995	0.1000	0.0631	0.0631	0.0501	0.0398	0.0316	0.0251	0.0251	0.0200
Max	1.0000	0.5012	0.2512	0.1995	0.1585	0.1259	0.1000	0.1000	0.1000	0.1000	0.1000
Bayesian Elastic Net		0.2864	0.0669	0.0254	0.1764	0.1086	0.1086	0.1086	0.0412	0.1764	0.2864

Note: For Elastic Net and Adaptive Elastic Net Regression, the table shows the median minimum and maximum value of lambda from 100 repetitions of the cross-validation procedure