ธนาคารแห่งประเทศไทย
BANK OF THAILAND

An economic application of machine learning:

Nowcasting Thai exports using global financial market data and time-lag lasso

PIER Exchange

Nov. 17, 2016

Thammarak Moenjak

What is machine learning?

Wikipedia (Nov 2016)

- "Evolved from the study of **pattern recognition** and **computational learning** theory in **artificial intelligence**, machine learning explores the study and construction of algorithms that can learn from and **make predictions on data**"

Motivation (1)

- Why nowcast Thai exports?
  - An important component of Thai GDP
  - Gives us ideas about the prevailing state of the economy
  - Test if machine learning might be useful in our context
    - Varian, 2014 and Einav and Leven, 2013 advocate the application of machine learning in economics
    - But of course:
      - Economics needs conherent interpretations (Most machine learning focus on predictive accuracy)
      - Most machine learning techniques currently available are for cross-sectional data

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Motivation (2)

Why use global financial market data to forecast exports?

– Domain knowledge

- Microeconomics (1st principles)
  - Demand (e.g. income, substitution, & wealth effects) and supply
- Macroeconomics and Finance
  - The yield curve is well documented to be a good predictor of the economy* (Use trading partners' and Thai gov't curves)
  - Stock markets partly represents financial wealth and shown to affect consumption**
  - NEER can provide a measure of substitution effects

* For example, Estrella and Trubin (2006), Estrella and Hadouvelis (1991)
**Juster et al. (2005), Dynan and Maki (2000), Porteba (2000)

Motivation (3)

Why use global financial market data to forecast exports (cont'd)?

- Information content
  - Prices reflect what people perceive about the future, with P&L actually taken into considerations
- Timeliness
  - Available on a near real-time basis

Motivation (4)

- Why use machine learning?
  - A large number of potential predictors relative to the number of observations
    - In this analysis: 148 predictors, but 136 (monthly) observations
  - Optimized for prediction accuracy
    - Not for the single covariate causal effects
  - Interpretability
    - v. traditional benchmark (ARIMA)
    - Of the particular technique used (Time-lag Lasso)

Global financial market data (1)

Monthly average, June 2005-Sep 2016

From Bloomberg

- The slope of the government bond yield curve (27 trading partners)*
- Stock market index (48 trading partners)
- Market conditions (Bcom, WTI)**

From BOT

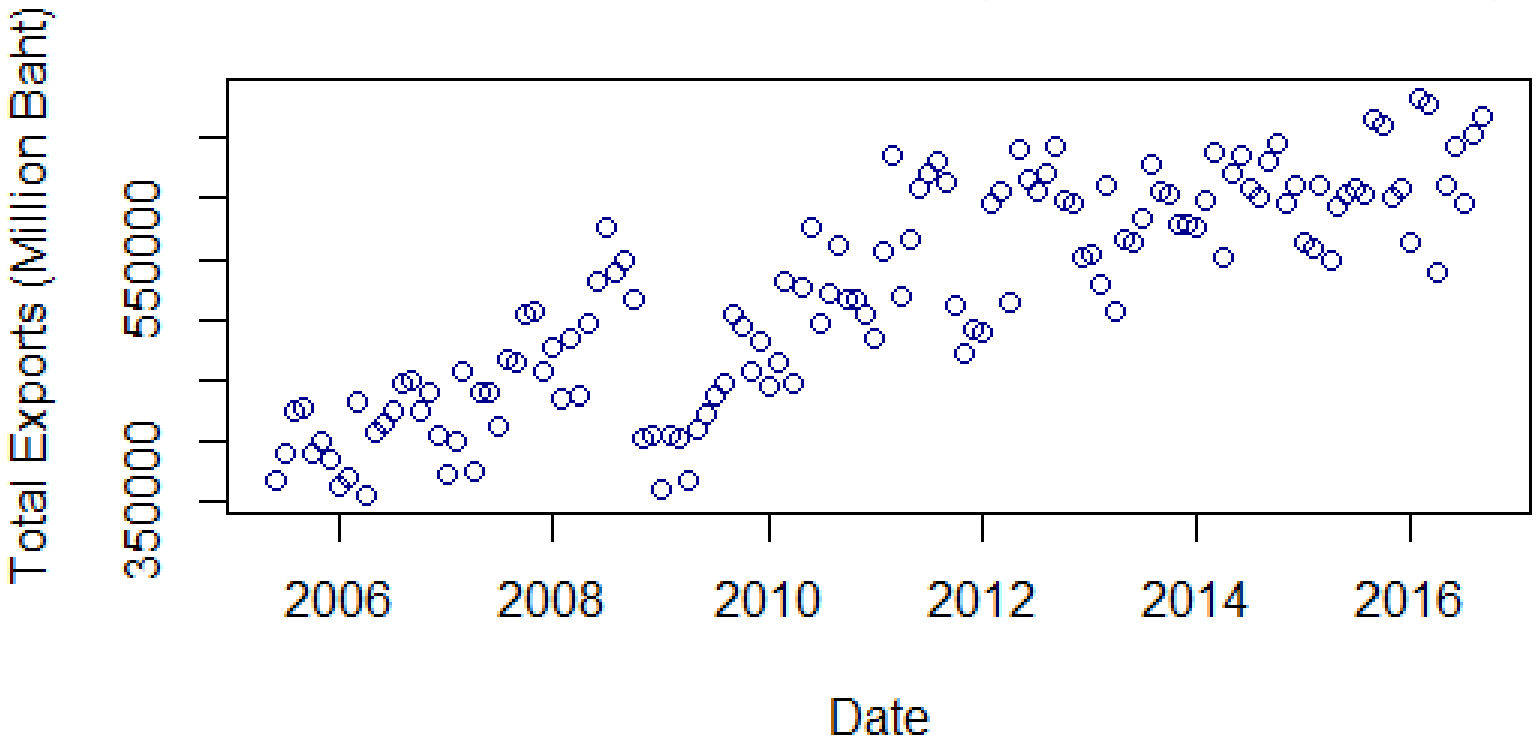
- NEER

Altogether: 74 variables

*2_10 slope, Except China, 2_7, Bloomberg Fair Value
** Bloomberg Commodity Index & West-Texas Index (Oil)

Figure 1:

Thai Exports (June 2005 - Sep 2016)



Augmented Dickey-Fuller Test

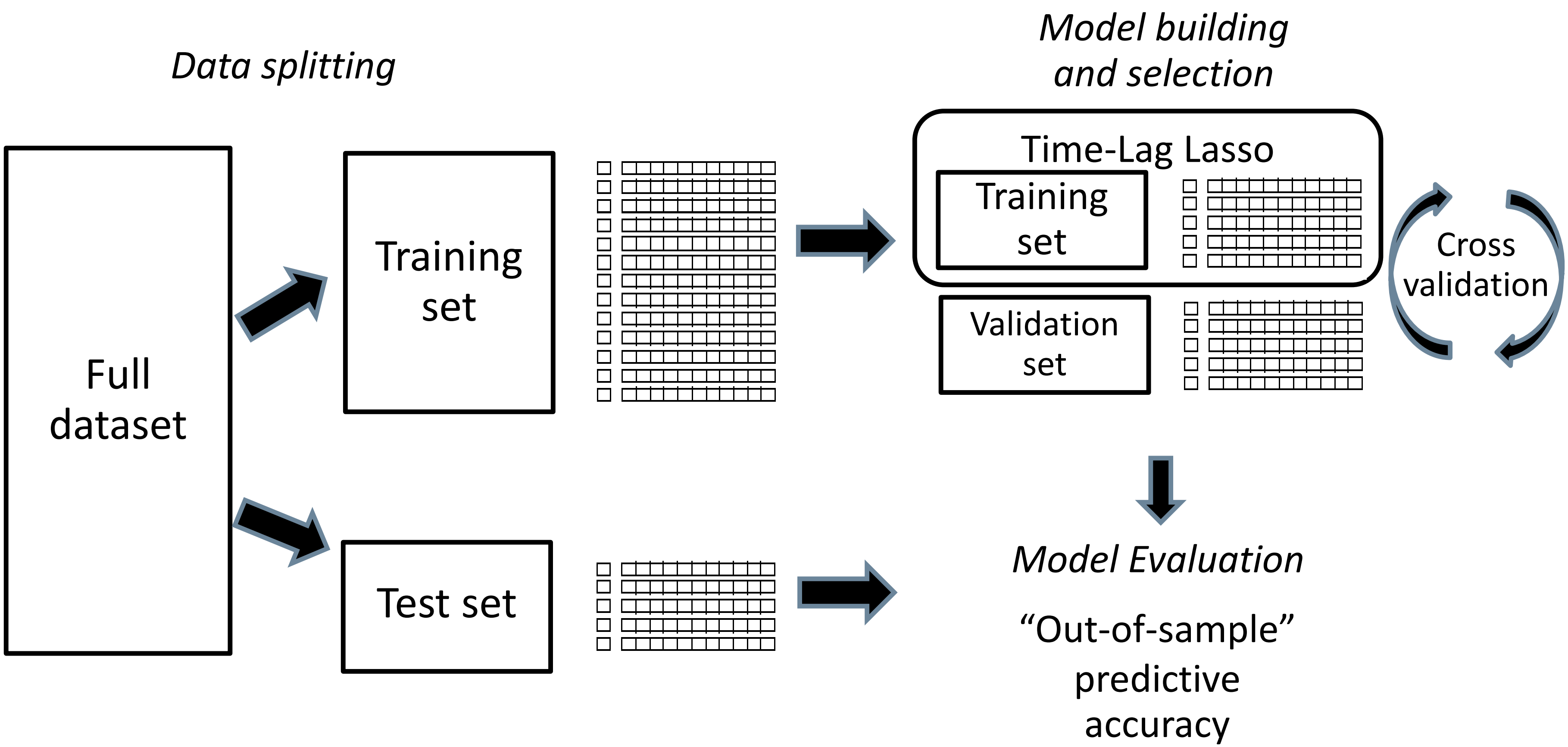
data: exportsx$Exports

Dickey-Fuller = -4.1309, Lag order = 5, p-value = 0.01

alternative hypothesis: stationary

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Overview of the Methodology



ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Thai exports:
Training and Test sets

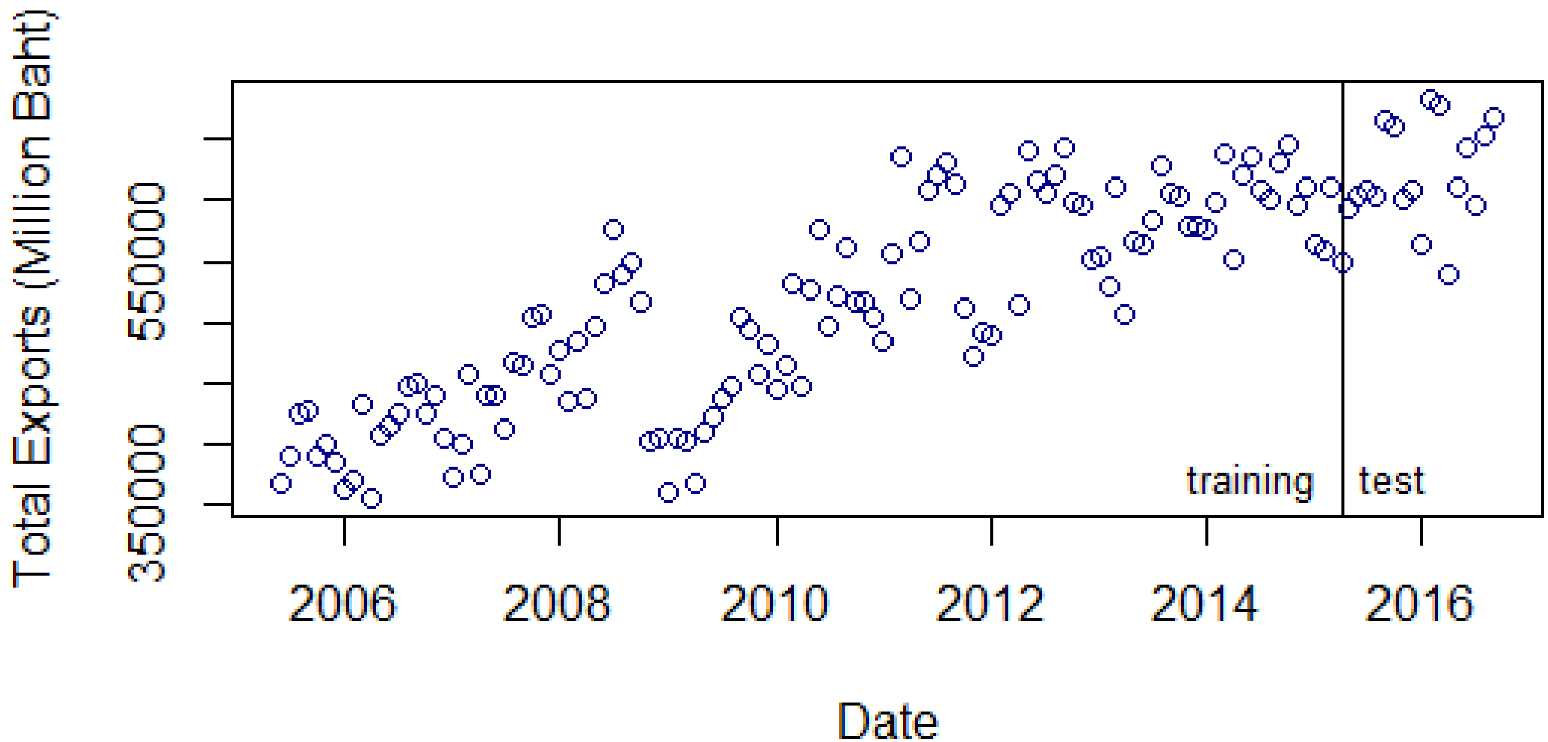
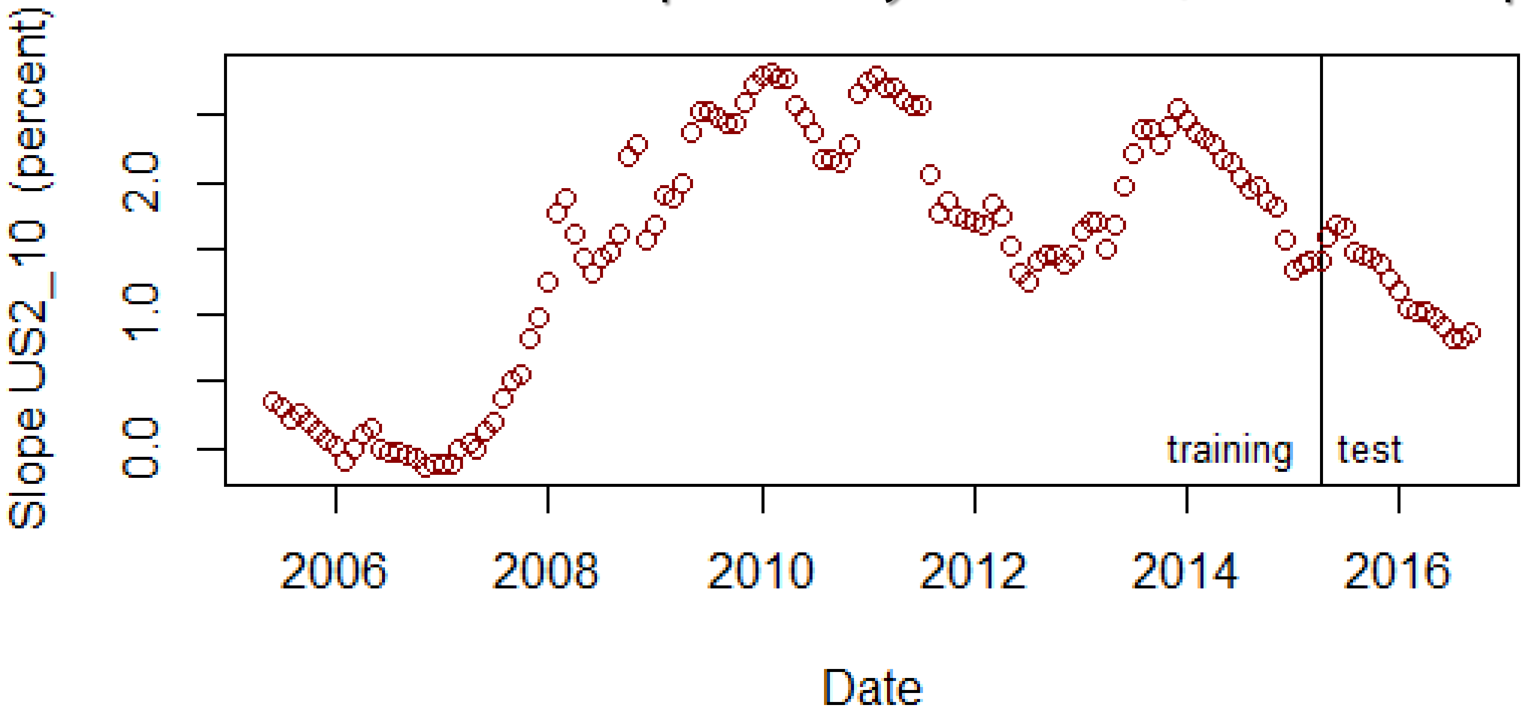


Figure 2.1:
Slope of US yield curve (June 2005-Sep 2016)

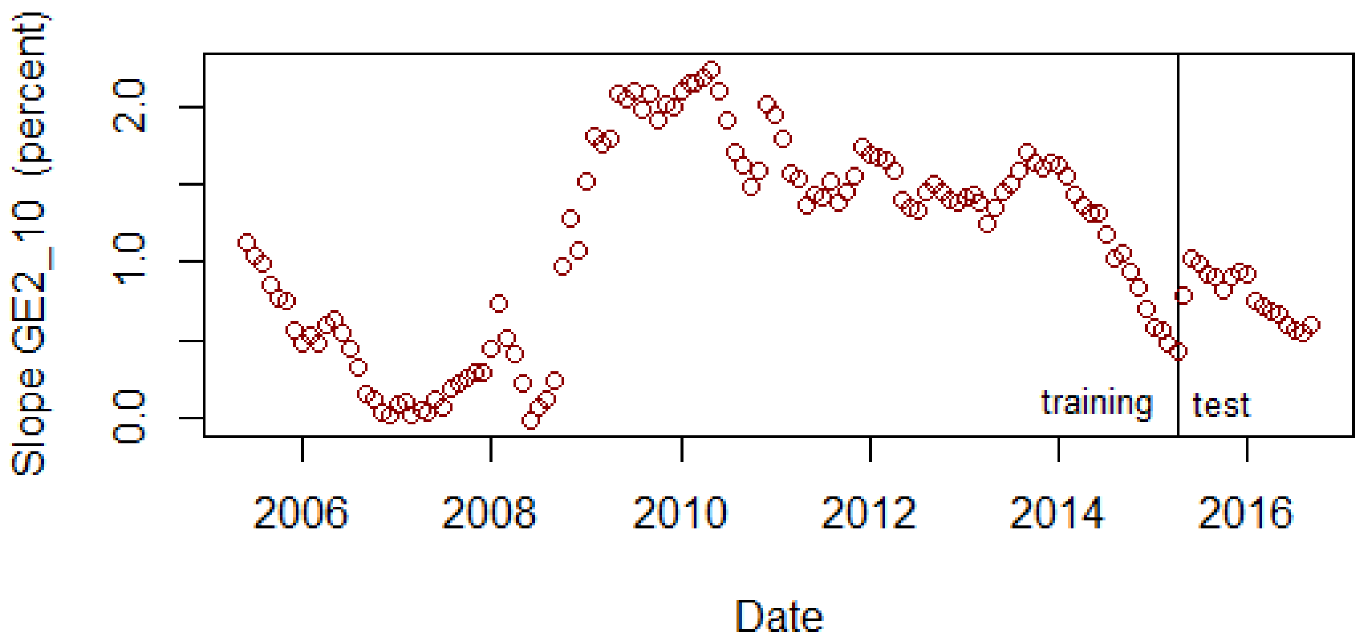


Augmented Dickey-Fuller Test
data: exportsx.new$UST2_10
Dickey-Fuller = -1.079, Lag order = 5, p-value = 0.9226
alternative hypothesis: stationary

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Figure 2.2:

Slope of German yield curve (June 2005-Sep 2016)

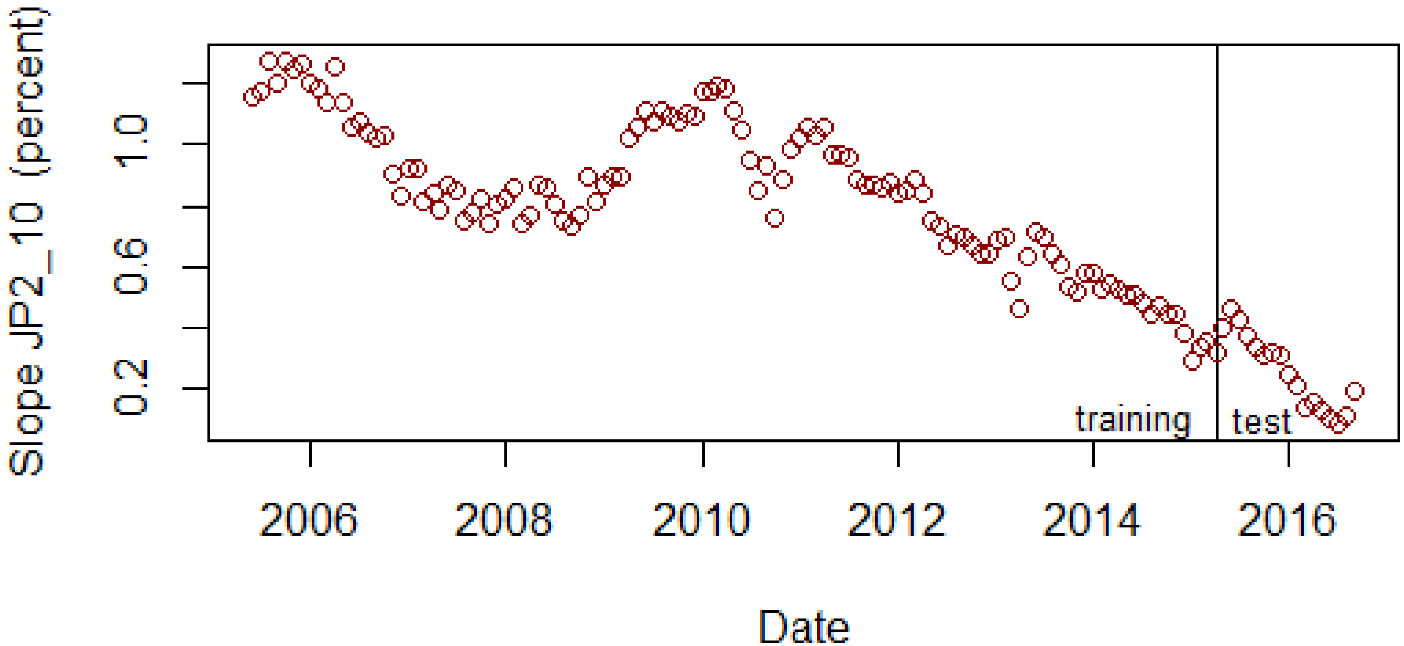


Augmented Dickey-Fuller
Test data: exportsx.new$GDBR2_10
Dickey-Fuller = -1.1909, Lag order = 5, p-value = 0.905
alternative hypothesis: stationary

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

2.3:

16)

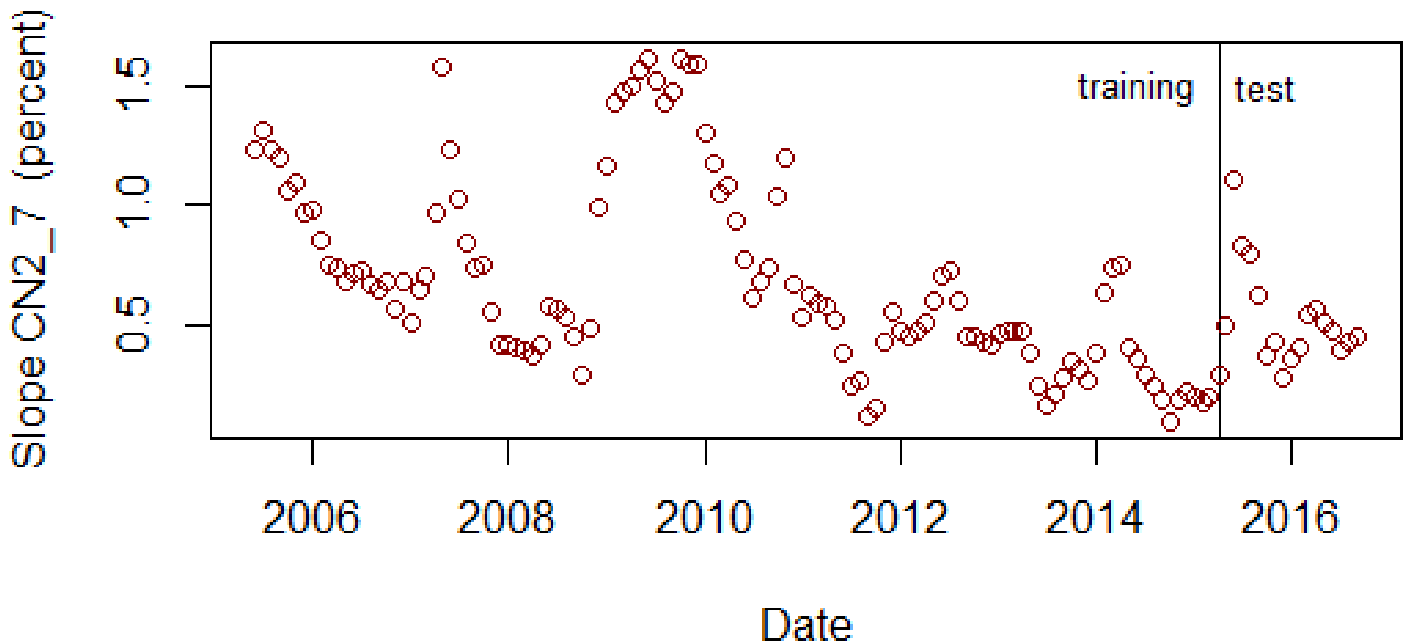


Augmented Dickey-Fuller Test
data: exportsx.new$JGB2_10
Dickey-Fuller = -1.6664, Lag order = 5, p-value = 0.7152
alternative hypothesis: stationary

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Figure 2.4:

)16)



Augmented Dickey-Fuller Test
data: exportsx.new$CNY2_7
Dickey-Fuller = -3.2046, Lag order = 5, p-value = 0.09013
alternative hypothesis: stationary

Figure 2.5:

Slope of Indonesia yield curve (June 2005-Sep 2016)

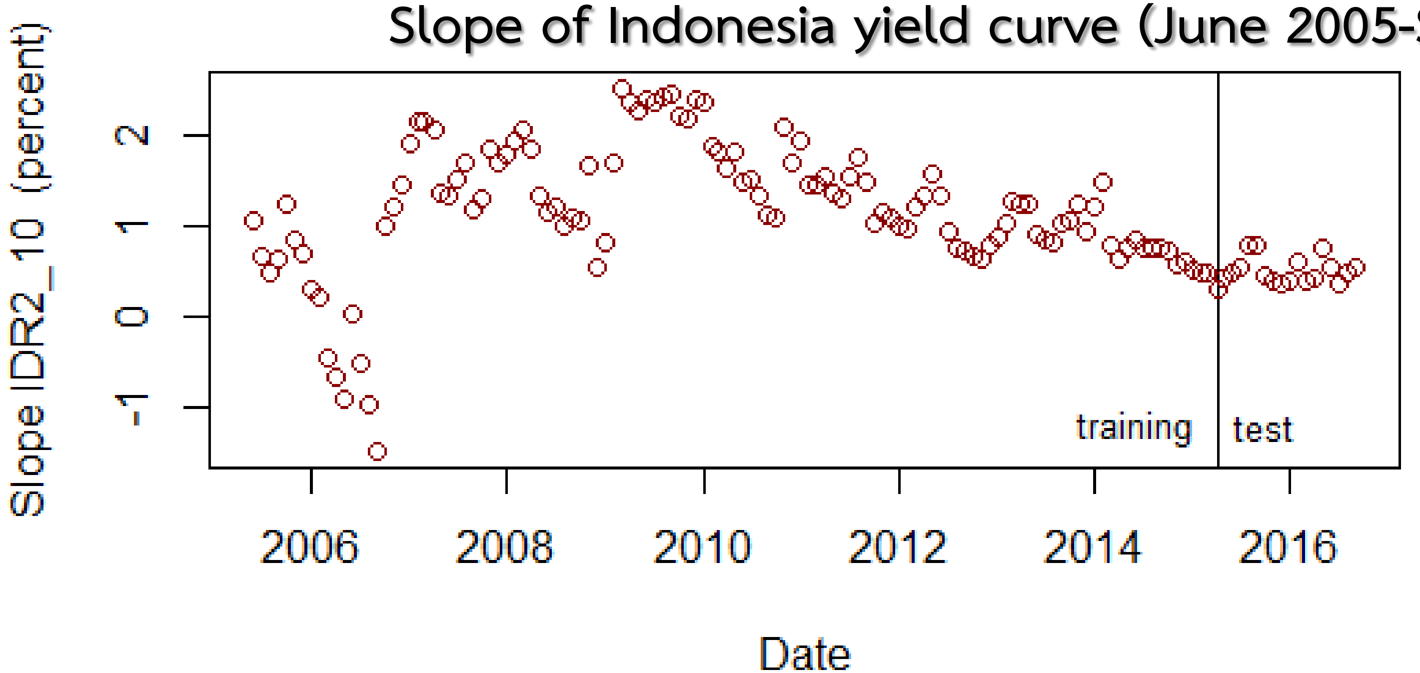
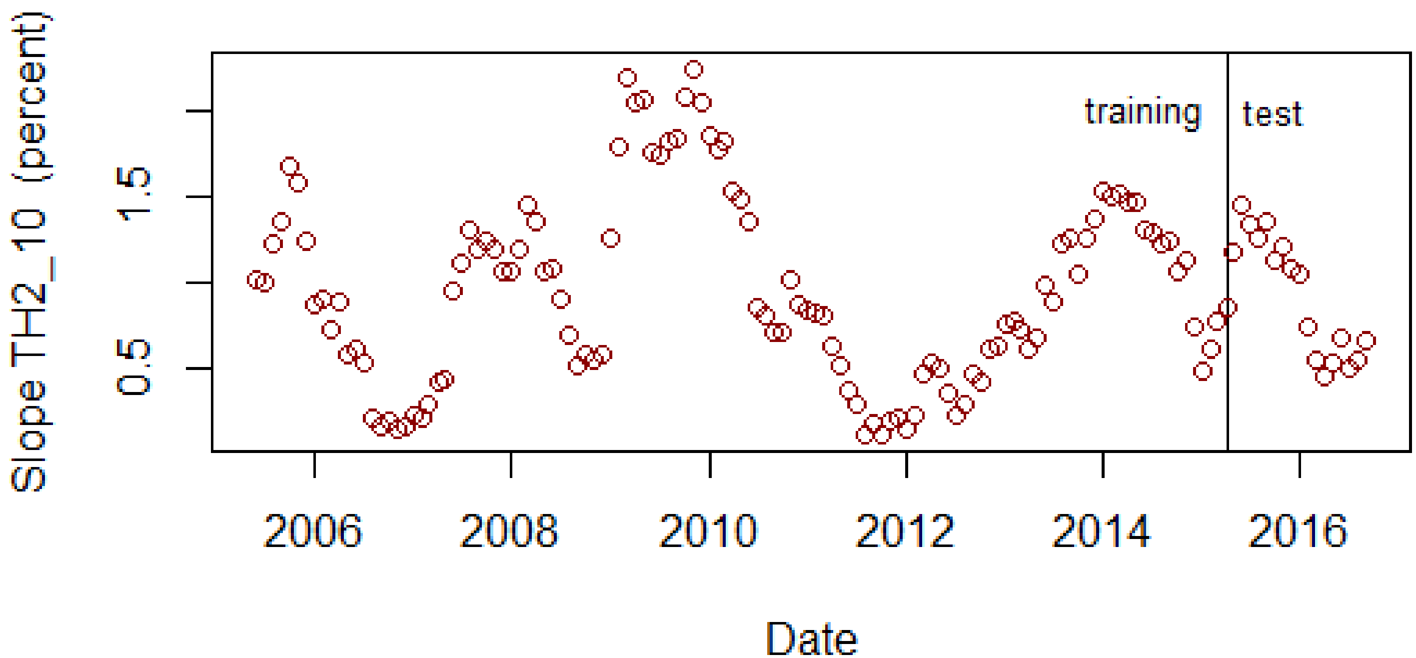


Figure 2.6:

Slope of Thailand yield curve (June 2005-Sep 2016)



ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Figure 2.7:
Bloomberg Commodity Index
(June 2005-Sep 2016)

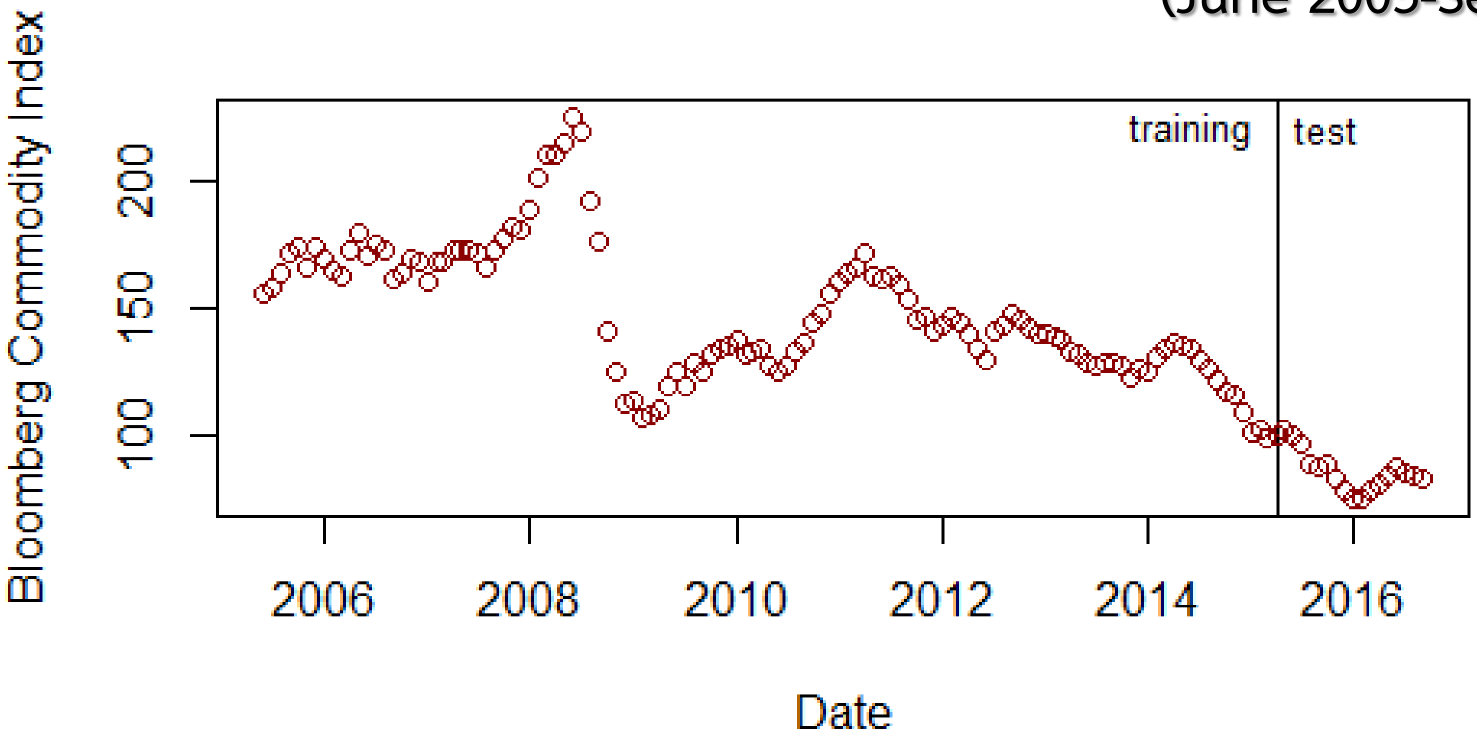


Figure 2.8:
NEER
(June 2005-Sep 2016)

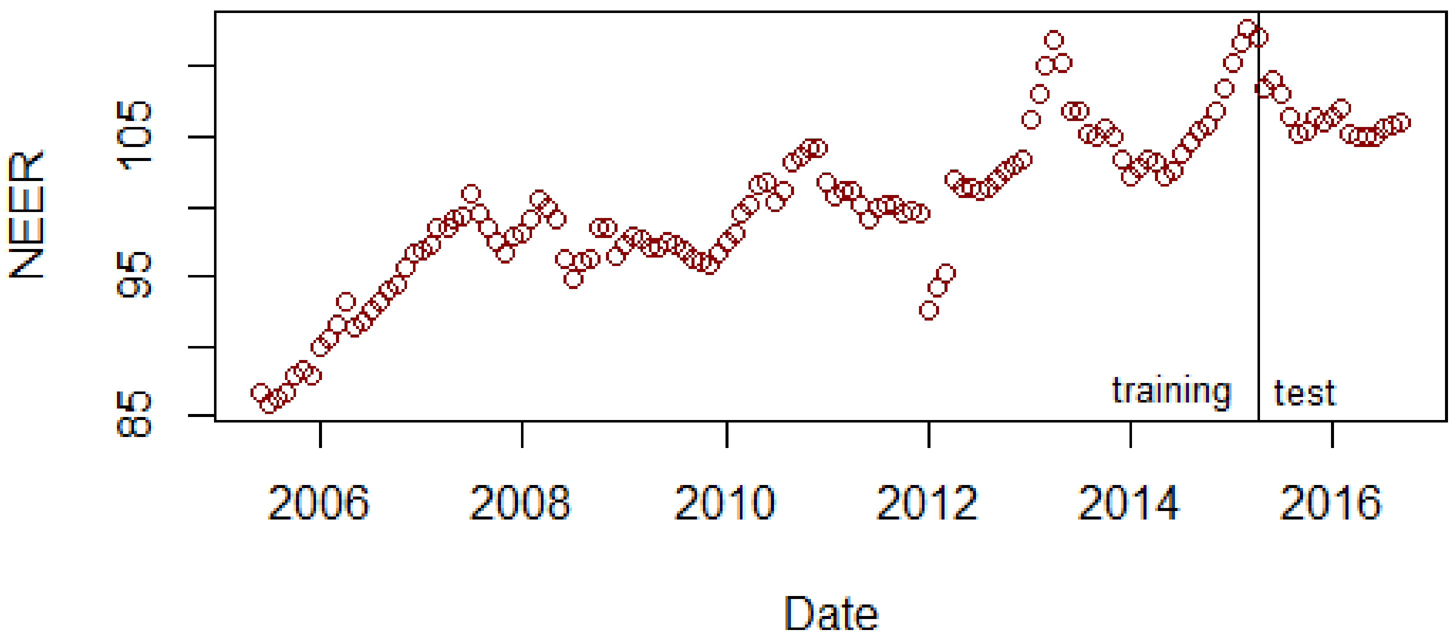
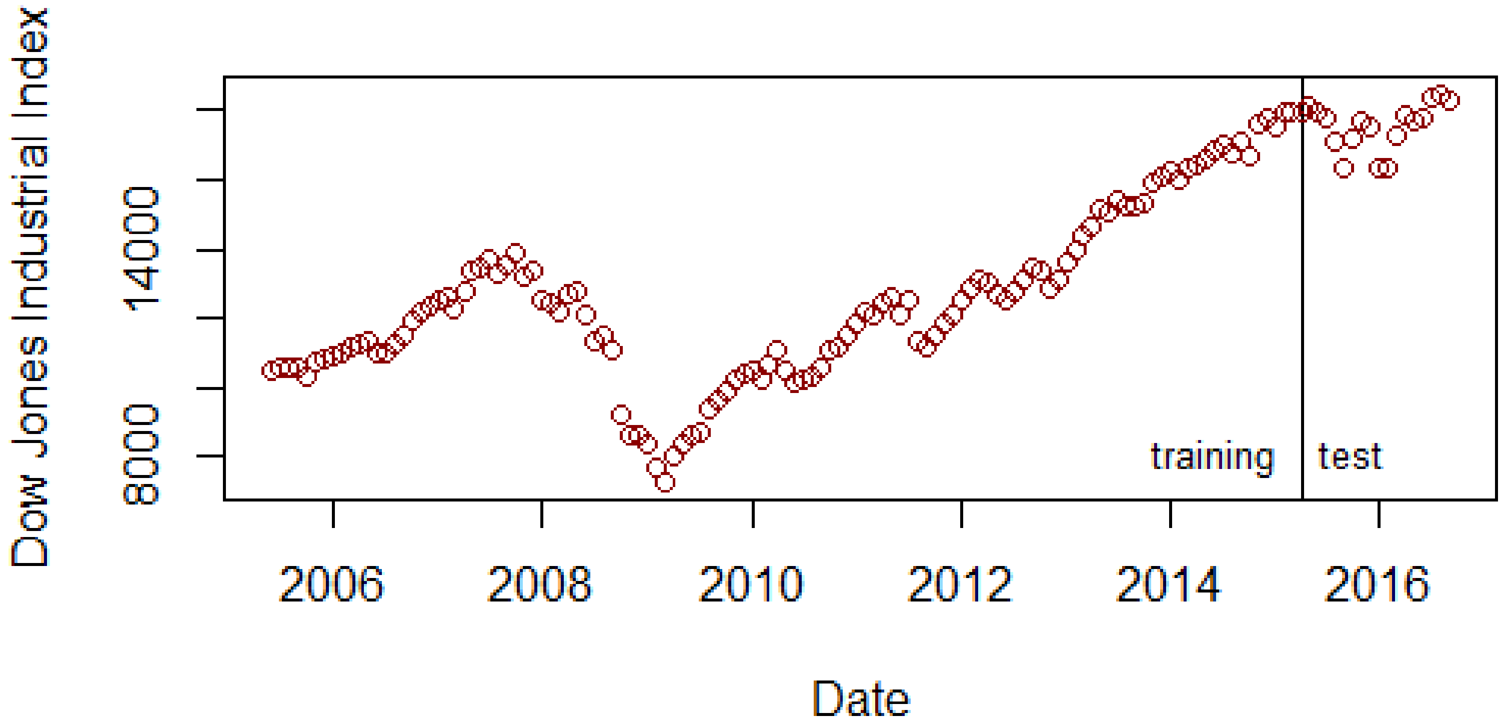
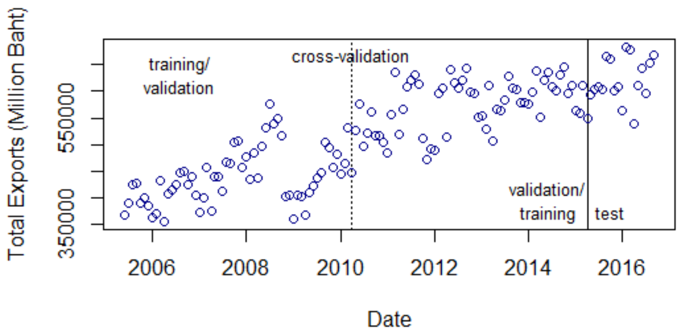
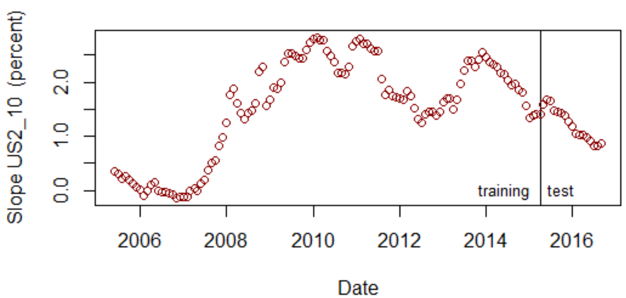
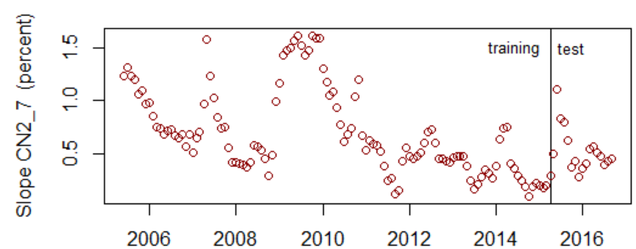


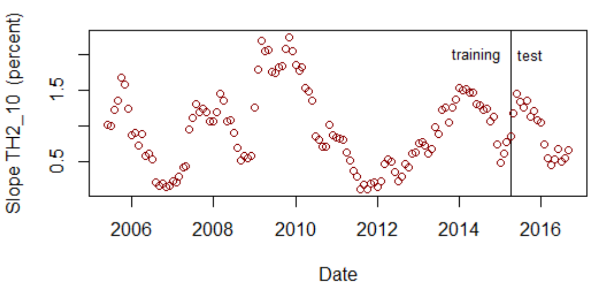
Figure 3.1:
Dow Jones Industrial Average
(June 2005-Sep 2016)

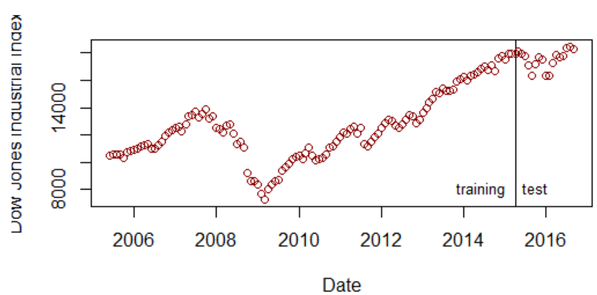


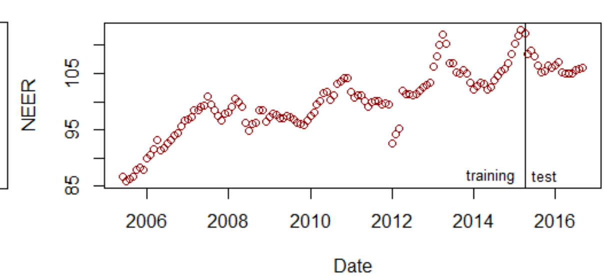

Total Exports (Million Baht)
350000
550000
training/
validation
cross-validation
validation/
training
test
2006
2008
2010
2012
2014
2016
Date


Slope US2_10 (percent)
0.0
1.0
2.0
training
test
2006
2008
2010
2012
2014
2016
Date


Slope CN2_7 (percent)
0.5
1.0
1.5
training
test
2006
2008
2010
2012
2014
2016


Slope TH2_10 (percent)
0.5
1.5
training
test
2006
2008
2010
2012
2014
2016
Date


8000
14000
training
test
2006
2008
2010
2012
2014
2016
Date


NEER
85
95
105
training
test
2006
2008
2010
2012
2014
2016
Date

Model building

The framework: The Lasso (Tibshirani, 1996)

(Least absolute shrinkage and selection operator)

Consider the usual linear regression framework:

$$y_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_{ij} + \epsilon_i \qquad (1)$$

with $E(\epsilon_i) = 0$ and $Var(\epsilon_i) = \sigma^2$ (2)

minimize $\left\{ \frac{1}{2} \sum_{i=1}^{N} (y_i - \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_{ij})^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\}$ (3)

Model building

The framework: The Lasso (Tibshirani 1996)

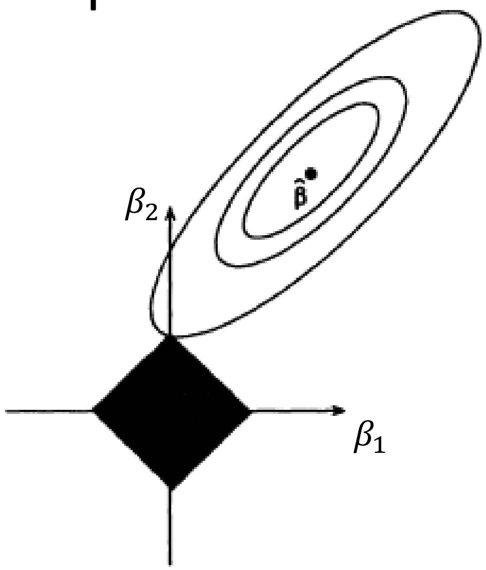
(Least absolute shrinkage and selection operator)

Or equivalently, minimize $\{RSS + \lambda \sum_{j=1}^{p} | \beta_j |\}$ (4)

Or, minimize$_{\beta}$ $\left\{ \sum_{i=1}^{N} (y_i - \beta_0 + \sum_{j=1}^{p} x_{ij} \beta_{ij})^2 \right\}$ s.t. $\sum_{j=1}^{p} | \beta_j | \leq s$ (5)

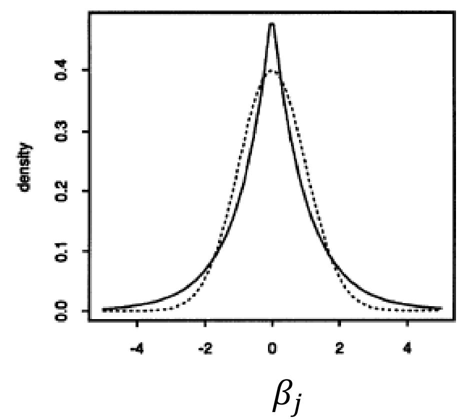
How does the Lasso perform variable selection?

Geometric interpretation



The ellipses are the contours of the RSS.
The solid triangle are the constraint regions $|\beta_1| + |\beta_1| \le s$

Probabilistic interpretation



The Lasso is the posterior mode for β under a double-exponential (Laplace) prior.
The solid line respresents a Laplace distribution.
The dotted line represents a normal distribution

Source: Tibshirani (1996)

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Time-Lag Lasso (1)

(Suo and Tibshirani, 2015)

Ordered Lasso:

$$\text{minimize} \quad \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(y_i - \beta_0 + \sum_{j=1}^{p} x_{ij} \beta_{ij} \right)^2 + \lambda \sum_{j=1}^{p} | \beta_j | \right\} \qquad (6)$$

Subject to $|\beta_1| \geq |\beta_2| \geq \ldots \geq |\beta_p|$

This makes sense in problems where some natural order exists among the predictors. But this problem is not convex, so modify the approach.

Write each $\beta_j = \beta_j^+ - \beta_j^-$,

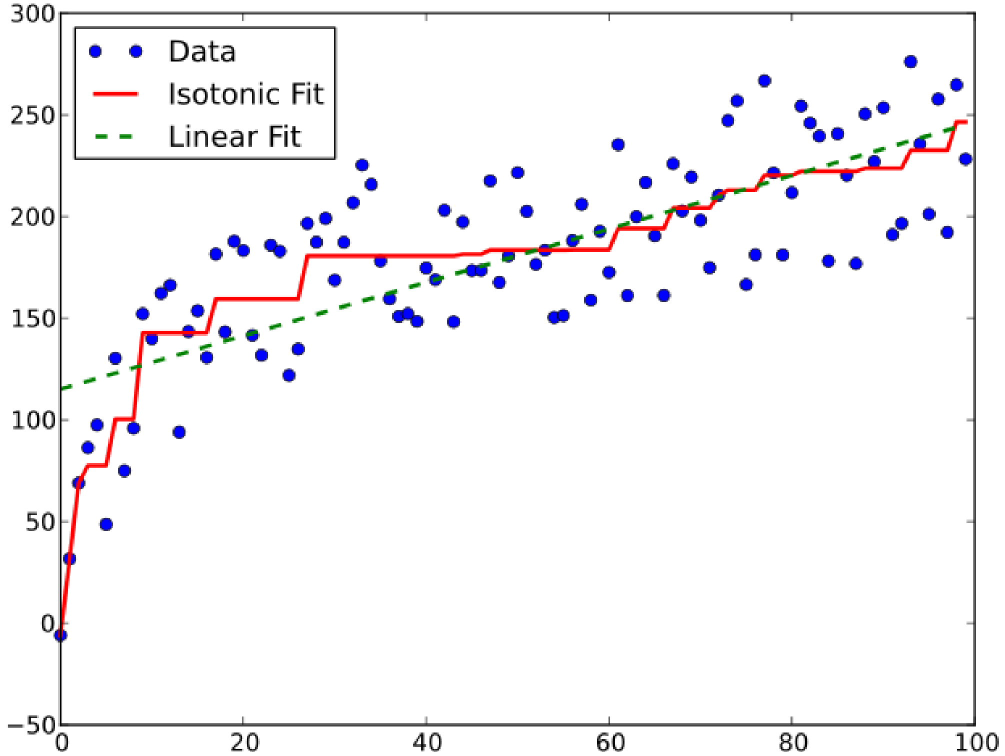
with $\beta_j^+, \beta_j^- \geq 0$,

$\beta_1^+ \geq \beta_2^+ \geq \ldots. \geq \beta_p^+ \geq 0$, and $\beta_1^- \geq \beta_2^- \geq \ldots. \geq \beta_p^- \geq 0$.

This could be solved by Pool Adjacent Violators Algorithm (PAVA).

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

PAVA is often used to solve isotonic regression problem



Isotonic regression: Fitting a free-form line of regression, whereby the fitted line has to be non-decreasing everywhere, and lie as close to the observations as possible.

Time-Lag Lasso (2)

(Suo and Tibshirani, 2015)

Time-Lag Lasso (rolling prediction)

The data has the form $\{y_{t,} x_{t1}, \dots, x_{tp}\}$ for *t = 1,2,…N* observations

The model has the form

$$y_t = \beta_0 + \sum_{j=1}^{p} \sum_{k=1}^{K} x_{t-k,j} \ \beta_{kj} + \epsilon_i \qquad (7)$$

with $E(\epsilon_i) = 0$ and $Var(\epsilon_i) = \sigma^2$,
$y_{t,}$ being the value of the outcome variable we want to predict at time *t.*
The value $x_{t-k,j}$ is the measurement of predictor *j* at time-lag k from the current time *t*.

Time-Lag Lasso (3)

Suo and Tibshirani (2015)

Write each $\beta_{kj} = \beta_{kj}^{+} - \beta_{kj}^{-}$ and propose the ff. problem

$$\text{minimize} \quad \{\frac{1}{2}\sum_{i=1}^{N}(y_t - \hat{y}_t)^2 + \lambda \sum_{j=1}^{p}\sum_{k=1}^{K}(\beta_{kj}^{+} + \beta_{kj}^{-})\} \tag{8}$$

Subject to $\beta_{1j}^{+} \geq \beta_{2j}^{+} \geq \ldots \geq \beta_{Kj}^{+} \geq 0$ and $\beta_{1j}^{-} \geq \beta_{2j}^{-} \geq \ldots \geq \beta_{Kj}^{-} \geq 0, \forall_j$

And solve by building a matrix **Z** of size $N \times (K_p)$**,** with K columns for each predictor. Each row of matrix **Z** has the form

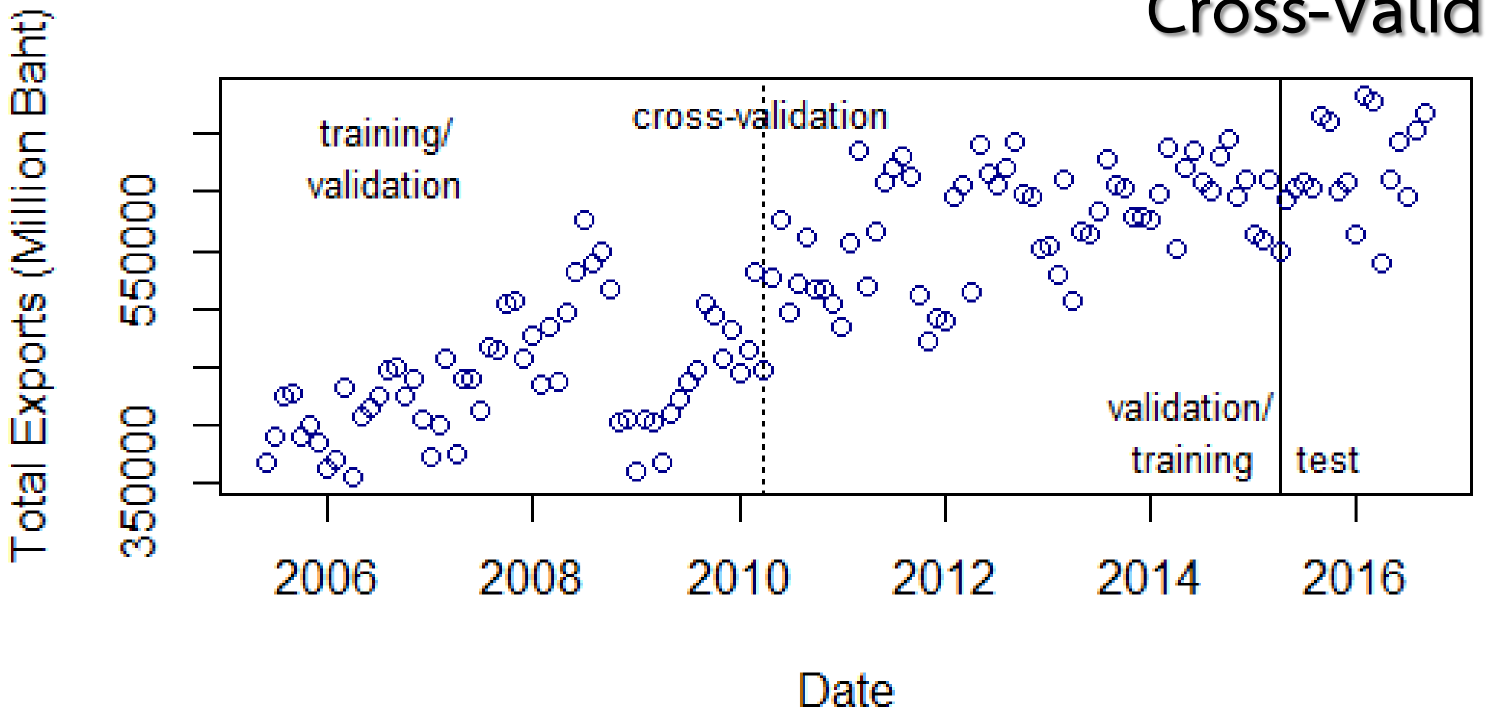
$$\{x_{t-1,1}, x_{t-2,1}, \ldots, x_{t-K,1} | x_{t-1,2}, x_{t-2,2}, \ldots, x_{t-K,2} | \ \ldots \ | \ x_{t-1,p}, x_{t-2,p}, x_{t-K,p}\}.$$

Each block corresponds to a predictor, lagged 1,2,..,K time units. There are N such rows. Augment each predictor $x_{t-k,j}$ with $x_{t-k,j}^{*} = -x_{t-k,j}$. Then apply block coordinate descent.

Model Selection:

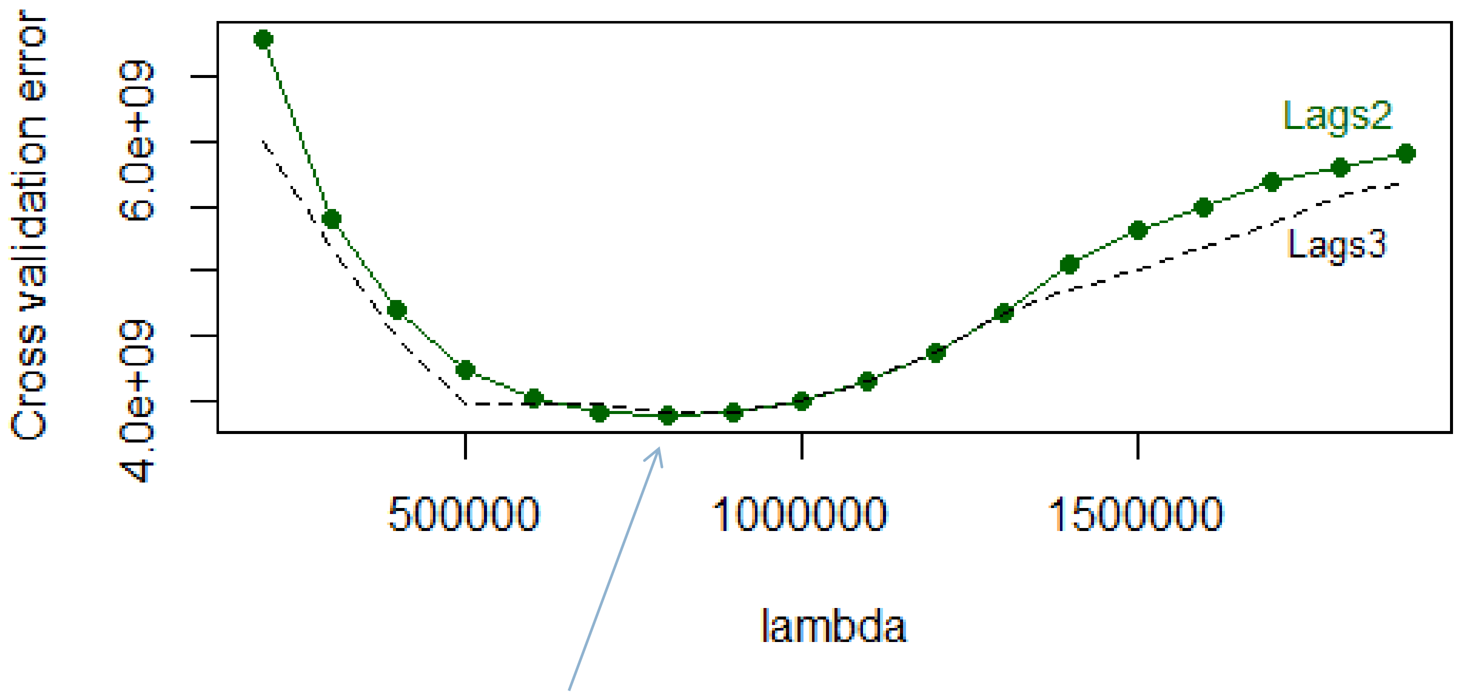
The tradeoff between bias and variance

Cross-Validation



ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Cross-validation and model selection



- Two hyper-parameters to tune: λ and k *(no. of lags)*
- Select the values of λ and k that yield the lowest cross validation error,
- Here cross-validation suggests a model with 8e+05, k = 2.

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Coefficients of the selected model

- (+) for advanced economies: US, NZ
- (-) for emerging market economies: CN, INR

-> Reflecting "Risk-on" & "Risk-off" conditions?

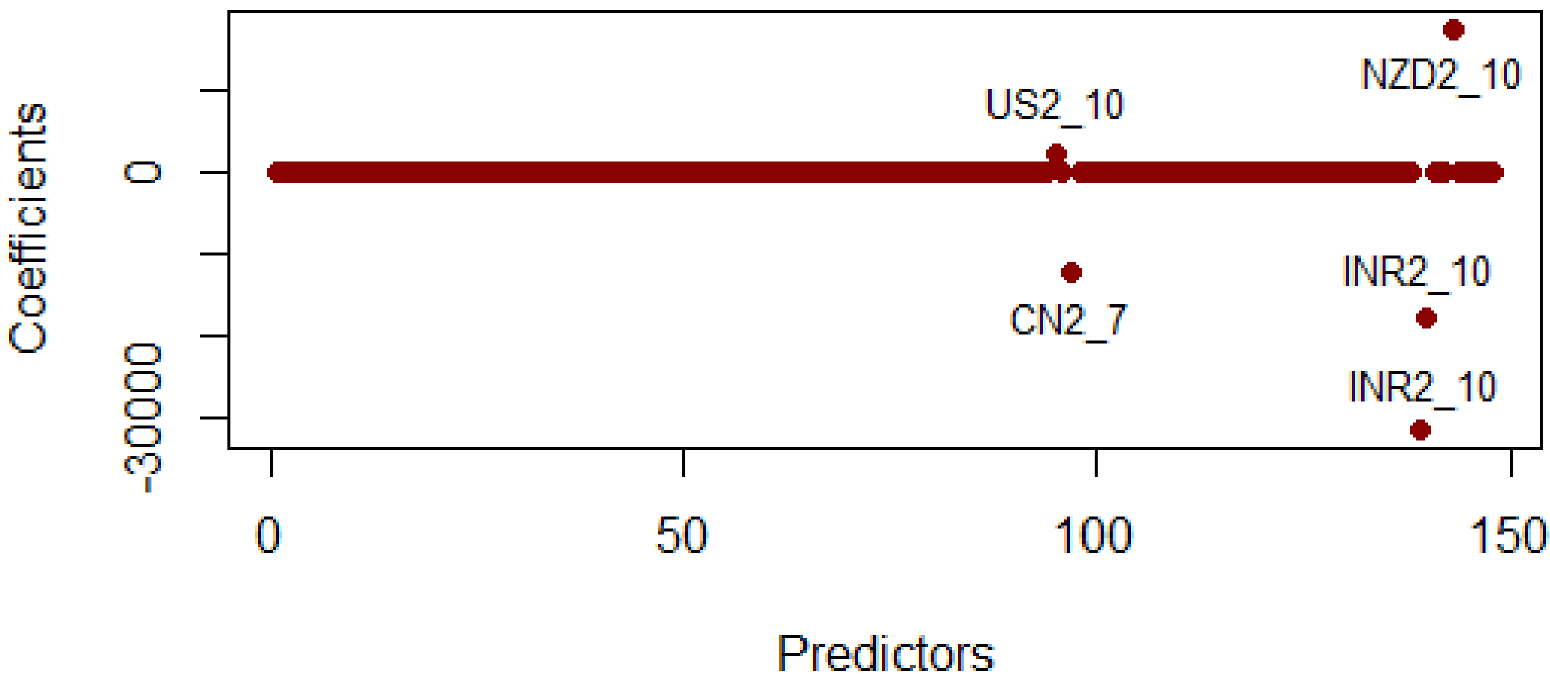


Figure 3.2:
Slope of India yield curve (June 2005-Sep 2016)

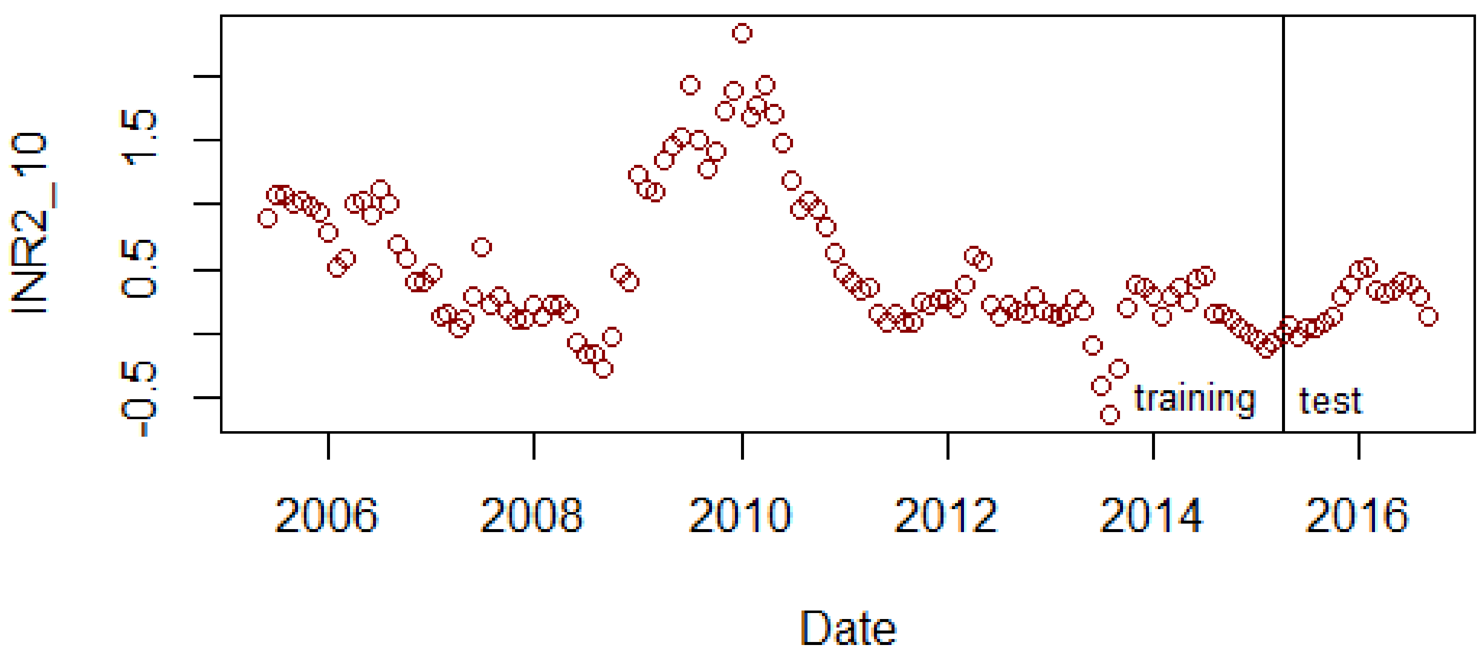
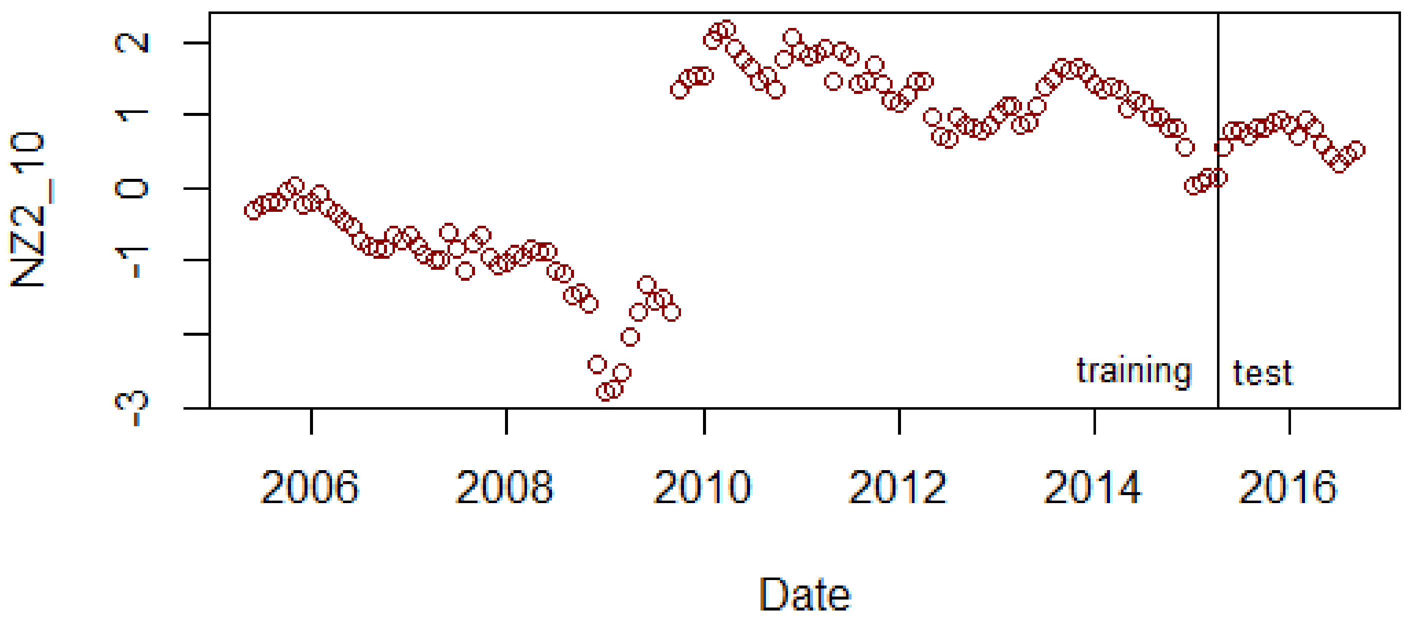
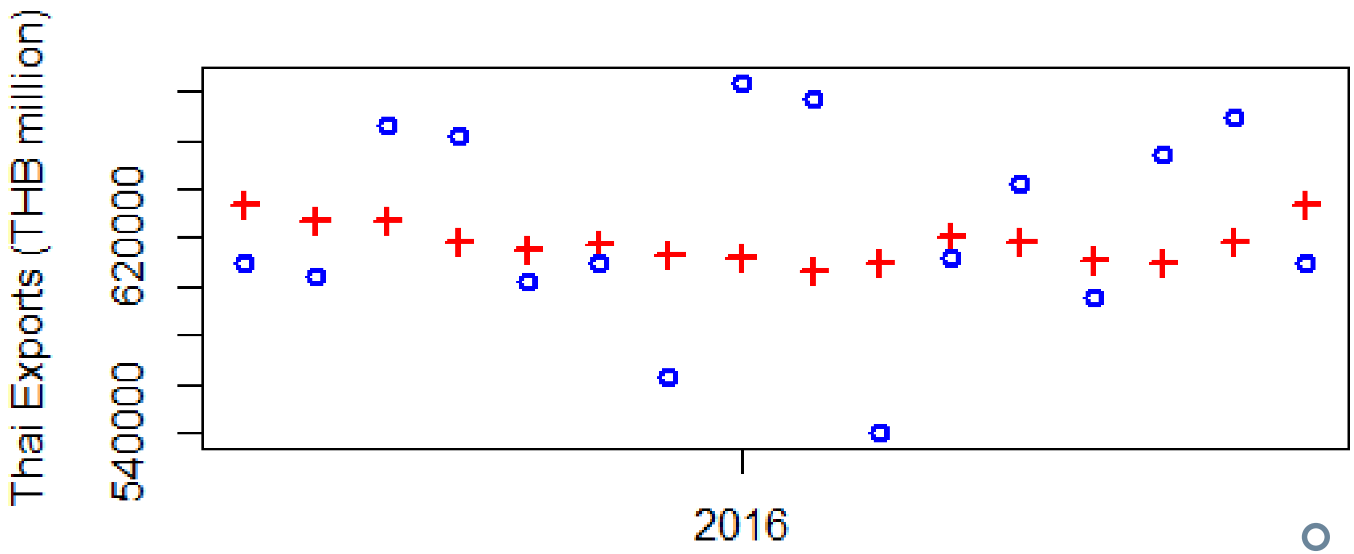


Figure 3.3:

Slope of New Zealand yield curve (June 2005-Sep 2016)



Out-of-sample (Test set)
Time-Lag Lasso v. Actual Thai Exports
(Jun 2015 to Sep 2016)

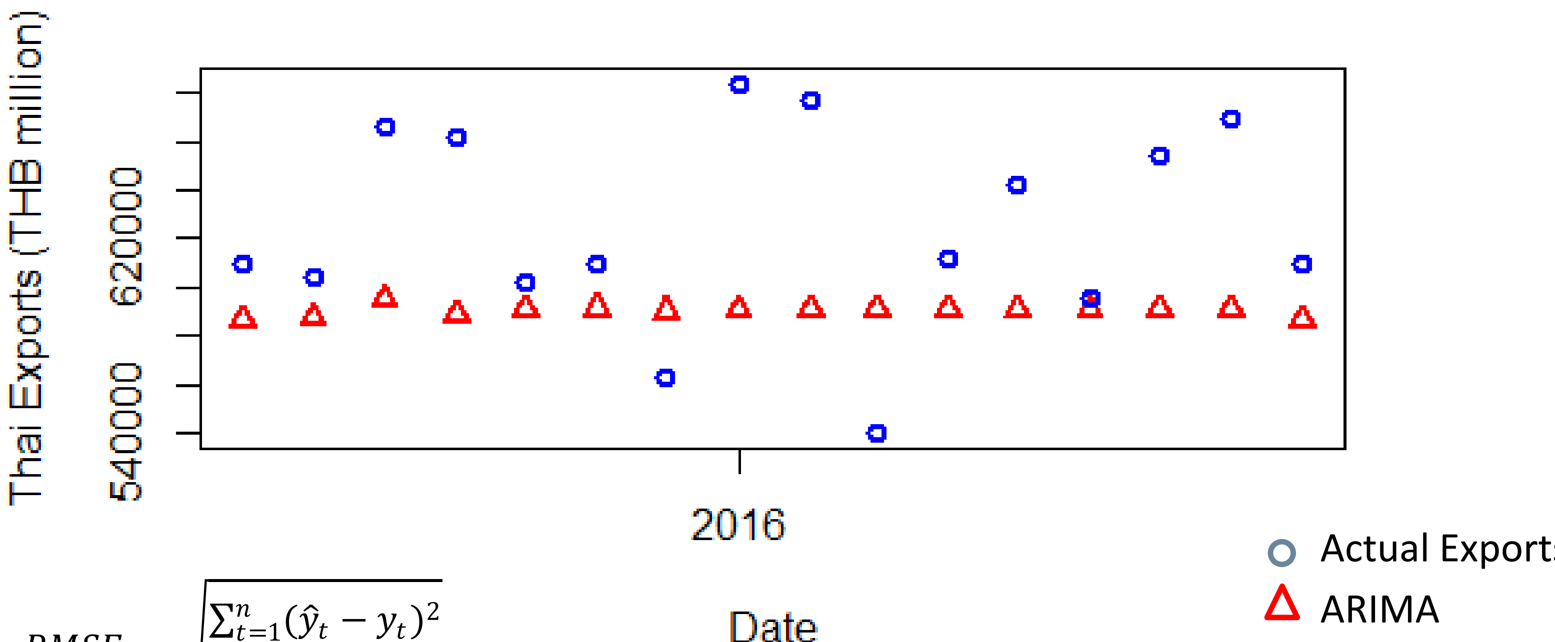


$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}}$$

$$CV(RMSE) = \frac{RMSE}{\bar{y}}$$

Time-Lag Lasso
Coefficient of variation of RMSE = 0.077

Out-of-sample (Test set)
ARIMA v. Actual Thai Exports
(Jun 2015 to Sep 2016)



$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}}$$

$$CV(RMSE) = \frac{RMSE}{\bar{y}}$$

ARIMA (2,1,1)

Coefficient of variation of RMSE = 0.097

Key takeaways (1)

- There are strong theoretical and empirical grounds for the use of global financial market data as predictors of Thai exports
- Time-Lag Lasso enables us to deal with time-series data *and* a large number of predictors
- The results seem to be make economic sense, while predictive accuracy is competitive with the traditional benchmark (ARIMA)

ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Key takeaways (2)

- Key machine learning ingredients have a good potential for policy applications
  - Regularization
  - Cross-Validation

- Caveats
  - This is quite different from traditional statistics with focuses on hypothesis testing
  - Coherent economic interpretations remain the most essential, choose your technique wisely
  - Also, relationships can change
  - And, as with *all* models, Lucas critique applies (esp. if we base our policy decisions on the models)