Is your nationality a secret?

–Evidence from 2016 ESS data and regularization method

Zihan Zhang, Zhaokun Zhang

April 20, 2019

Abstract

In machine learning, Lasso, Ridge regression and ealstic net are most common and widespread methods of regularization. In this artilce, we combine these methods with an intersting expriment–predict your nationality based on some characteristics and cognition. Through the analysis we hope to have a beter understanding of regularization and model selection.

1. Introduction

1.1 Lasso

In statistics and machine learning, Lasso(least absolute shrinkage and selection operator) is a regression analysis method which performs both variable selection and regularization. It aims at improving the prediction accuracy and interpretability of regression models by altering the model fitting process to select only a subset of the provided covariates for use in the final model rather than using all of them. And it's basic form is:

$$\hat{\beta}^L = argmin\{\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p} \mid \beta_j \mid\}$$

where $\lambda \geq 0$ is a tuning parameter that controls the strength of penalty term, which has the effect of shrinking the estimates towards zero. Equivalently,

$$\hat{\beta}^L = argmin\{\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2\}$$

subject to $\sum_{j=1}^{p} \mid \beta_j \mid \leq C$

The Lasso shrinks the coefficient estimates towards zero, but in this case, the penalty has the effect of forcing some of the coefficient estimates to be exactly equal to zero when the tuning parameter(λ) is sufficiently large. Thus, we say that Lasso yields sparse models that involve only a subset of variables.

1.2 Ridge Regression

Ridge regression is one of most commonly used method of regularization of illposed problems. Compared to Lasso regression, ridge regression cannot make some coefficients be zero when shrinking. In addition, β_0 is not penalized, that is, we don't shrink it. The estimates from the ridge method are defined by

$$\hat{\beta}^R = argmin\{\sum_{i=1}^{N}(y_i - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}\beta_j^2\}$$

Note that ridge regression estimates are not scale equivalent. x_{ij} need to be standardized before applying it. Then the formula becomes:

$$\hat{\beta}^R = argmin\{\sum_{i=1}^{N}(\tilde{y_i} - \sum_{j=1}^{p}\beta_j \tilde{x}_{ij})^2 + \lambda\sum_{j=1}^{p} \mid\mid \beta_j \mid\mid_2^2\}$$

where $|| \, . \, ||_2$ denotes the l_2 norm, $\tilde{y_i} = y_i - \bar{y}$, $\tilde{x}_i = [\tilde{x}_{i1}, ..., \tilde{x}_{ip}]'$

1.3 Elastic net

In the fitting of linear or logistic models, the elastic net is a regularized regression method that linearly combines the l_1 and l_2 penalties of the Lasso and ridge methods. It overcomes the limitations of Lasso like highly correlated variables or high dimensional data. The formula is :

$$\Omega(\beta) = \alpha \mid\mid \beta \mid\mid_1 + (1-\alpha) \mid\mid \beta \mid\mid_2^2$$

The elastic net shrinks together the coefficients of correlated features like ridge regression, and it also encourages sparse solutions like Lasso.

2. Empricial Analysis

2.1 Data and variables

In order to have a further knowledge of these method of regularization, we use data from *Europen Social Survey(ESS)*, which contains a wide array of survey data from European countries, revealing the cognition and attitudes to a couple of social status and problems of residents in various regions. All those cognition and attitudes were reported in some specific themes including **media and social trust, politics, gender, social demographics, human values, walfare attitudes, religion and social attitude towards climate change**. Each theme comprises a series of questions that contribute to manifest precise descriptions of personal preference of individuals. Besides, the scale of measurement are usually described within 0-10, representing the magnitude that interviewee agree with this question.

The latest data was the 8th round one in the year of 2016 with 534 variables and 44,387 observations from 23 countries initially. From our points of view, due to the various characters of countries, residents in different countries may have different preferences on those themes. Hence, from the whole dataset, we reasonably selected 152 variables that may be helpful to distinguish their nationalities. Meanwhile, we drew out 8812 individuals from five particular countries including the UK, France, Italy, Sweden and Russia because of their different geographic location and cultural background. For example, as our intuition, Englishmen are commonly recognized conservative while Russian may be more resistant to homosexuality, citizens in France and Italy suffer more from the refugees while the Swedish concerns more about their welfare provided by their country. The following figure indicates the location of countries involved in our classification.

Hence, we came up with an idea that predicting the nationalities of individuals with their personal information, based on the belief of citizens in various countries with different views on the same questions. With multinomial regression and methods of model selection and regularization, we can not only find out the best model for prediction, but also figure out the most important factors distinguishing people from various countries.

Besides, before running regressions, we cleaned the data and omitted some observations with missing values. Finally the dataset contained 2882 individuals from 5 countries.

```r
# install.packages("glmnet")
rm(list = ls())
library(AER)
library(mlogit)
library(glmnet)
library(nnet)
```

```r
####### import data ############

ESS8_c5 <- read.table("ESS8_c5.csv",header=TRUE,sep=",",row.names="number")
ess <- na.omit(ESS8_c5)
```

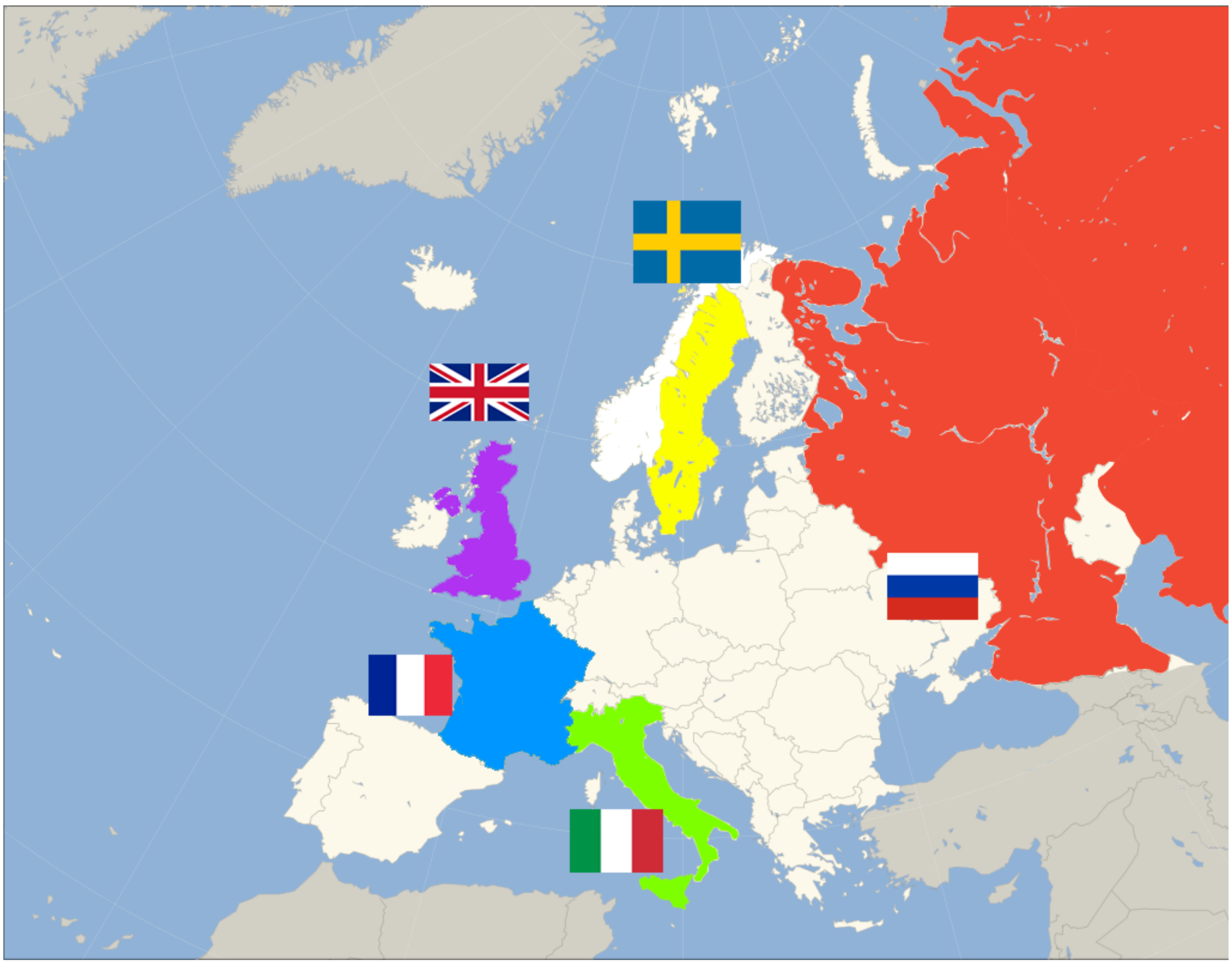


Figure 1: Euromap

```r
n=nrow(ess)
train <- sample(n,0.8*n)
ess.tr <- ess[train,] ## Training set
ess.test <- ess[-train,] ## Test set

x <- data.frame(ess.tr[,3:148])
y <- ess.tr$cntry
x1 <- as.matrix(x)
```

2.2 The model

In this part, we introduce the regression and regularization model we use. Before the regression, we divided our data set into two parts: training set used to fit the parameters of the model and test set used to provide an unbiased evaluation of a model fit on the training dataset. Firstly, we use multnomial logit model to do a simple regression. On the basis of the regression, we do a predict test in our test set to verify the model, and we also calculate the error of this model. The results are as follows:

```r
######## Multnomial Logit Regression #######
attach(ess.tr)
mlfit <- multinom(cntry ~ . ,data = data.frame(ess.tr[,2:148]))
mlfit.yhat <- predict(mlfit, ess.test)
```

```r
t1 <- table(mlfit.yhat,ess.test$cntry,dnn = c("predicted","true"))
t1
```

```
##          true
## predicted  FR  GB  IT  RU  SE
##        FR 170   8   0   0   6
##        GB   6 148   3   5   6
##        IT   2   4  51   0   1
##        RU   1   2   0  25   0
##        SE   4  10   2   1 122
```

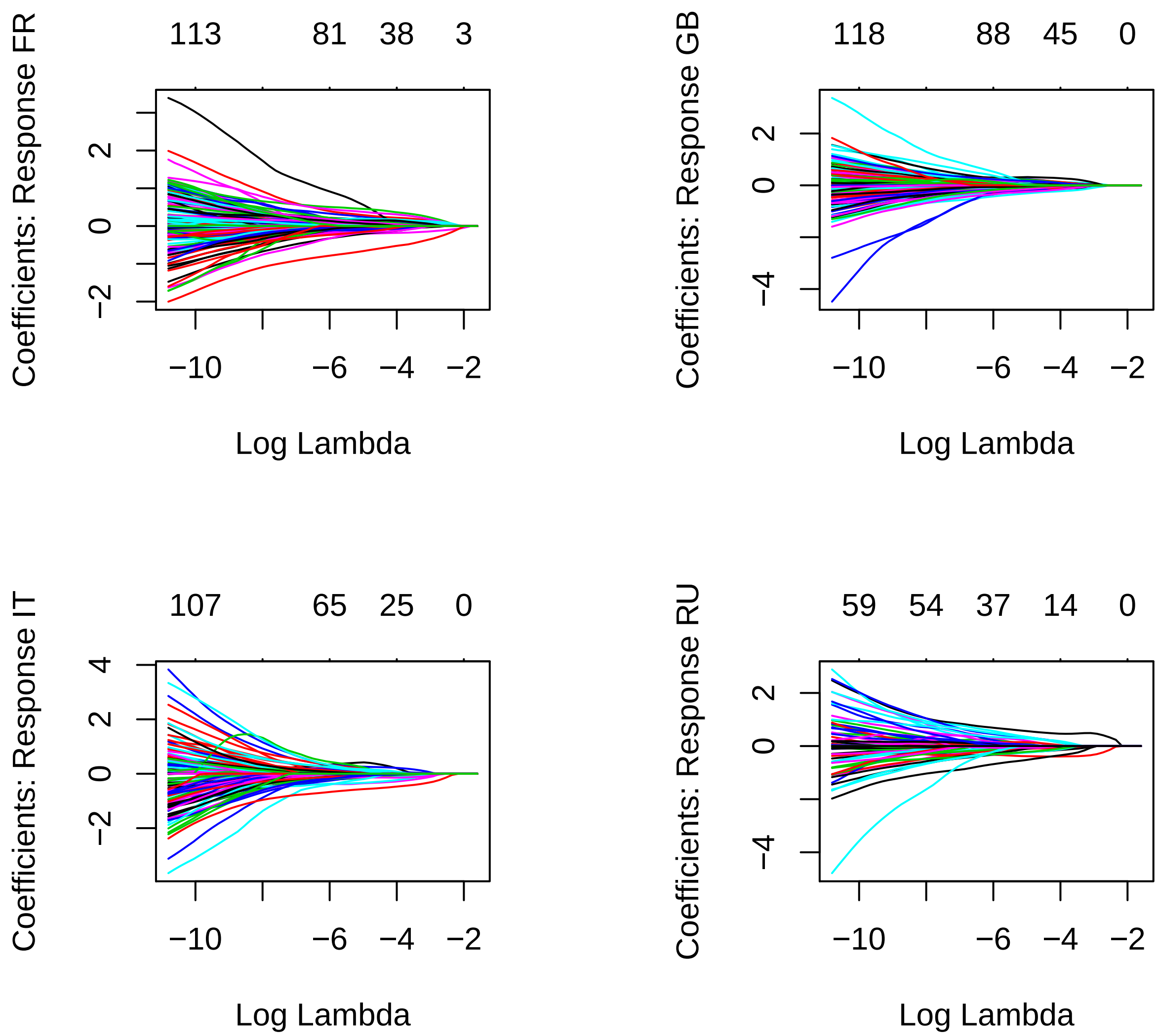
```r
pred_accuracy_multi <- 1 - sum(diag(t1))/sum(t1)
pred_accuracy_multi
```

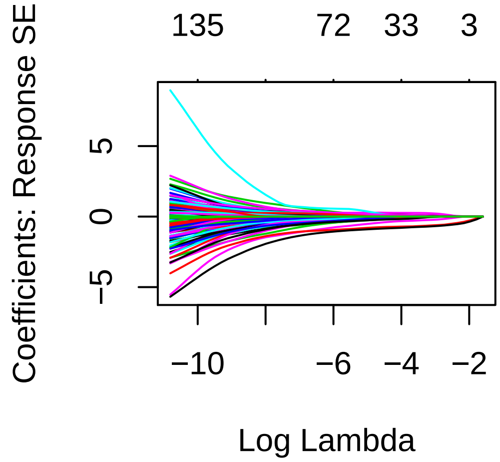
```
## [1] 0.1057192
```

```r
detach(ess.tr)
```

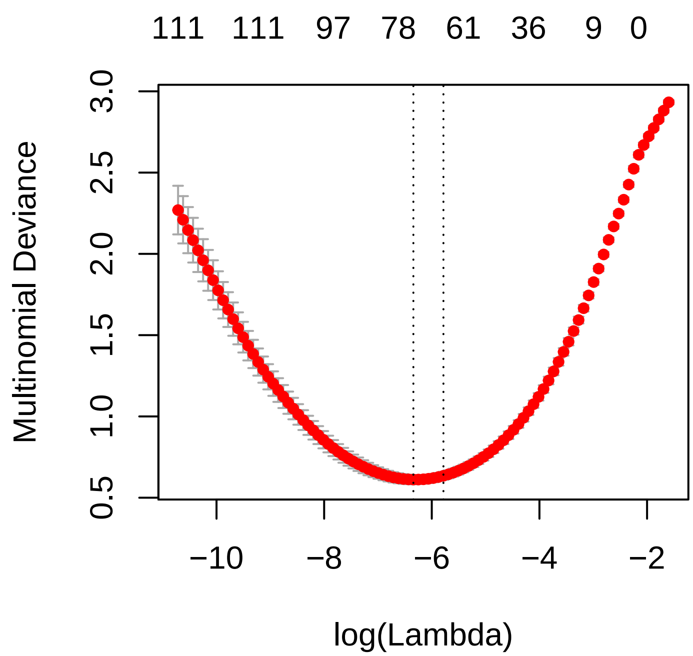
After using multinomial logit regression model, we then use Lasso regression method to select variables as well as regularize it. Note that before applying the regression, the data is standardized in order to predict more accurately. In the following we only take out the Lasso figures of 5 different countries and the optimal λ which has the lowest error:

```r
########## Cross validation LASSO ###########
lassofit <- glmnet(x1,y,family = "multinomial",standardize = T,nlambda = 50,alpha = 1)
# coef(lassofit)
plot(lassofit,xvar = "lambda",label = F)
```


113 81 38 3
Coefficients: Response FR
−2 0 2
−10 −6 −4 −2
Log Lambda
118 88 45 0
Coefficients: Response GB
−4 0 2
−10 −6 −4 −2
Log Lambda
107 65 25 0
Coefficients: Response IT
−2 0 2 4
−10 −6 −4 −2
Log Lambda
59 54 37 14 0
Coefficients: Response RU
−4 0 2
−10 −6 −4 −2
Log Lambda




```
cvlassofit <- cv.glmnet(x1,y,family = "multinomial",standardize = T,alpha = 1)
coef(cvlassofit)
plot.cv.glmnet(cvlassofit,sign.lambda = 1)
plot(cvlassofit)
```




```
optimal_lambda <- cvlassofit$lambda.min
```

```
optimal_lambda
```

```
## [1] 0.001760996
```

From the output, the λ should be as above. Then we use optimal λ to fit the regression model, but during the process we **omit** the results showing coefficients of variables picked since it can occupy too much space.

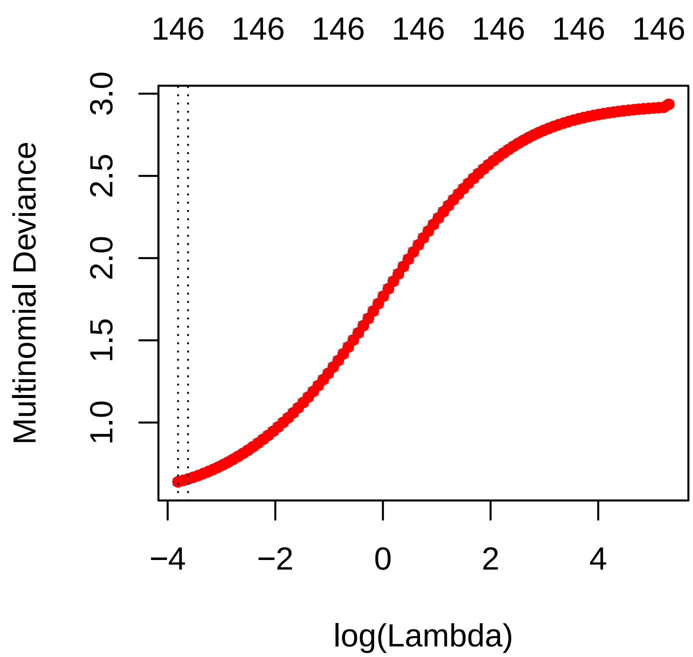
```
# Use the optimal lambda to fit the model
lassofit <- glmnet(x1,y,family = "multinomial",standardize = T,lambda = optimal_lambda,
                   alpha = 1)
coef(lassofit)
```

Furthermore, we use ridge regression and elastic net to regularize the model. All data are also standardized before applying the regression, and the coefficients are also omitted to save space. The figures and optimal λ are listed respectively.

```
######## Cross Validation RIDGE ###########
ridgefit <- cv.glmnet(x1,y,family = "multinomial",standardize = T, alpha = 0)
optimal_lambda_ridge <- ridgefit$lambda.min
optimal_lambda_ridge
```

```
## [1] 0.02222125
```

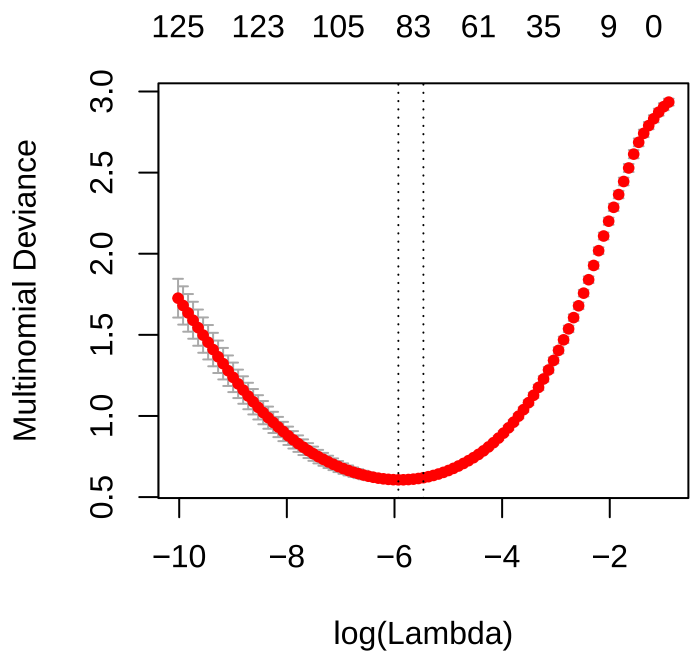
```
plot(ridgefit)
```




```
######## cross Validation Elastic net #######
enfit <- glmnet(x1,y,family = "multinomial",standardize = T,lambda = 0.2, alpha=0.5)
en_05 <- cv.glmnet(x1,y,family = "multinomial",standardize = T, alpha=0.5)
optimal_lambda_en <- en_05$lambda.min
optimal_lambda_en
```

```
## [1] 0.002664258
```

```r
plot(en_05)
```



2.3 Discussion of results

In this part we mainly focus on the results of prediction and classification error of different models in test set as well as comparing with original error by multinomial logistic regression which has already shown at the beginning.

```r
######## Lasso #######
lassofit.yhat <- predict(lassofit,newx = as.matrix(ess.test[,3:148]),s = optimal_lambda,
                         type = "class")
t2 <- table(lassofit.yhat,ess.test$cntry)
t2
```

```
##
## lassofit.yhat  FR  GB  IT  RU  SE
##            FR 173   9   1   0   3
##            GB   5 152   3   2   5
##            IT   2   2  50   0   0
##            RU   1   0   0  28   0
##            SE   2   9   2   1 127
```

```r
pred_accuracy_lasso <- 1 - sum(diag(t2))/sum(t2)
pred_accuracy_lasso
```

```
## [1] 0.08145581
```

```r
######## Ridge#######
ridgefit.yhat <- predict(ridgefit,newx = as.matrix(ess.test[,3:148]),s = optimal_lambda_ridge,
                         type="class")
```

```r
t3 <- table(ridgefit.yhat,ess.test$cntry)
t3
```

```
##
## ridgefit.yhat  FR  GB  IT  RU  SE
##            FR 171   5   1   1   3
##            GB   6 159   1   1   4
##            IT   1   2  51   0   1
##            RU   1   0   0  28   0
##            SE   4   6   3   1 127
```

```r
pred_accuracy_ridge <- 1 - sum(diag(t3))/sum(t3)
pred_accuracy_ridge
```

```
## [1] 0.07105719
```

```r
######## Elastic Net #######
enfit.yhat <- predict(en_05,newx = as.matrix(ess.test[,3:148]),s = optimal_lambda_en,
                      type="class")
t4 <- table(enfit.yhat,ess.test$cntry)
t4
```

```
##
## enfit.yhat  FR  GB  IT  RU  SE
##         FR 172   8   1   0   3
##         GB   5 154   3   1   4
##         IT   2   2  51   0   1
##         RU   1   0   0  29   0
##         SE   3   8   1   1 127
```

```r
pred_accuracy_en <- 1-sum(diag(t4))/sum(t4)
pred_accuracy_en
```

```
## [1] 0.0762565
```

From the results, it's obvious that through regularization methods we can make the predicted error much smaller in test set. And the output of these three models is similar.Then we can get a conclusion that these methods of regularization do improve the prediction accuracy.

Meanwhile, it can also illustrate that the variables picked at the beginning do have an effect on the prediction of person's nationality. Although there are some variables exclude in Lasso regression, they still have impact on dependent variable in other methods.

3. Reference

[1] G. Tutz, W.Poessnecker.(2012): *"Variable Selection for the Multinomial Logit Model"*

[2] Wikipedia, *"Lasso (statistics)"*, https://en.wikipedia.org/wiki/Lasso_(statistics)

[3] Wikipedia, *"Elastic net regularization"*, https://en.wikipedia.org/wiki/Elastic_net_regularization

[4] Jiaming Mao, *"Model Selection and Regularization"*,https://jiamingmao.github.io/data-analysis/assets/Lectures/Model_Selection_and_Regularization.pdf

Appendix

Table 1: Interpretation of variables

Variables	Theme	Interpretation
cntry		Country
netusoft	Society	Internet use, how often
ppltrst	Society	Most people can be trusted or you can't be too careful
pplfair	Society	Most people try to take advantage of you, or try to be fair
pplhlp	Society	Most of the time people helpful or mostly looking out for themselves
polintr	Society	How interested in politics
psppsgva	Society	Political system allows people to have a say in what government does
actrolga	Society	Able to take active role in political group
psppipla	Society	Political system allows people to have influence on politics
cptppola	Society	Confident in own ability to participate in politics
trstprl	Society	Trust in country's parliament
trstlgl	Society	Trust in the legal system
trstplc	Society	Trust in the police
trstplt	Society	Trust in politicians
trstprt	Society	Trust in political parties
trstep	Society	Trust in the European Parliament
trstun	Society	Trust in the United Nations
lrscale	Politics	Placement on left right scale
stflife	Politics	How satisfied with life as a whole
stfeco	Politics	How satisfied with present state of economy in country
stfgov	Politics	How satisfied with the national government
stfdem	Politics	How satisfied with the way democracy works in country
stfedu	Politics	State of education in country nowadays
stfhlth	Politics	State of health services in country nowadays
gincdif	Politics	Government should reduce differences in income levels
mnrgtjb	Politics	Men should have more right to job than women when jobs are scarce
freehms	Politics	Gays and lesbians free to live life as they wish
hmsfmlsh	Politics	Ashamed if close family member gay or lesbian
hmsacld	Politics	Gay and lesbian couples right to adopt children
euftf	Politics	European Union: European unification go further or gone too far
imsmetn	Politics	Allow many/few immigrants of same race/ethnic group as majority
imdfetn	Politics	Allow many/few immigrants of different race/ethnic group from majority
impcntr	Politics	Allow many/few immigrants from poorer countries outside Europe
imbgeco	Politics	Immigration bad or good for country's economy
imueclt	Politics	Country's cultural life undermined or enriched by immigrants
imwbcnt	Politics	Immigrants make country worse or better place to live
happy	Personality	How happy are you
sclmeet	Personality	How often socially meet with friends, relatives or colleagues
inprdsc	Personality	How many people with whom you can discuss intimate and personal matters
sclact	Personality	Take part in social activities compared to others of same age
crmvct	Personality	Respondent or household member victim of burglary/assault last 5 years
aesfdrk	Personality	Feeling of safety of walking alone in local area after dark
health	Personality	Subjective general health
hlthhmp	Personality	Hampered in daily activities by illness/disability/infirmity/mental problem
atchctr	Personality	How emotionally attached to [country]
atcherp	Personality	How emotionally attached to Europe
rlgdgr	Personality	How religious are you
rlgatnd	Personality	How often attend religious services apart from special occasions
pray	Personality	How often pray apart from at religious services
dscrgrp	Personality	Member of a group discriminated against in this country
gvrfgap	Personality	Government should be generous judging applications for refugee status
rfgfrpc	Personality	Most refugee applicants not in real fear of persecution own countries

rfgbfml	Personality	Granted refugees should be entitled to bring close family members
eneffap	Energy	How likely to buy most energy efficient home appliance
rdcenr	Energy	How often do things to reduce energy use
cflsenr	Energy	How confident you could use less energy than now
elgcoal	Energy	How much electricity in [country] should be generated from coal
elgngas	Energy	How much electricity in [country] should be generated from natural gas
elghydr	Energy	How much electricity in [country] should be generated from hydroelectric power
elgnuc	Energy	How much electricity in [country] should be generated from nuclear power
elgsun	Energy	How much electricity in [country] should be generated from solar power
elgwind	Energy	How much electricity in [country] should be generated from wind power
elgbio	Energy	How much electricity in [country] should be generated from biomass energy
wrpwrct	Energy	How worried, power cuts
wrenexp	Energy	How worried, energy too expensive for many people
wrdpimp	Energy	How worried, [country] too dependent on energy imports
wrdpfos	Energy	How worried, [country] too dependent on fossil fuels
wrntdis	Climate	How worried, energy supply interrupted by natural disasters or extreme weather
wrinspw	Climate	How worried, energy supply interrupted by insufficient power generated
wrtcfl	Climate	How worried, energy supply interrupted by technical failures
wrtratc	Climate	How worried, energy supply interrupted by terrorist attacks
clmchng	Climate	Do you think world's climate is changing
clmthgt2	Climate	How much thought about climate change before today
wrclmch	Climate	How worried about climate change
ccgdbd	Climate	Climate change good or bad impact across world
lkredcc	Climate	Imagine large numbers of people limit energy use, how likely reduce climate chan
lklmten	Climate	How likely, large numbers of people limit energy use
gvsrdcc	Climate	How likely, governments enough countries take action to reduce climate change
ownrdcc	Climate	How likely, limiting own energy use reduce climate change
inctxff	Climate	Favour increase taxes on fossil fuels to reduce climate change
sbsrnen	Climate	Favour subsidise renewable energy to reduce climate change
banhhap	Climate	Favour ban sale of least energy efficient household appliances to reduce climate
dfincac	Living Standard	Large differences in income acceptable to reward talents and efforts
smdfslv	Living Standard	For fair society, differences in standard of living should be small
uemplwk	Living Standard	Of every 100 working age how many unemployed and looking for work
slvpens	Living Standard	Standard of living of pensioners
slvuemp	Living Standard	Standard of living of unemployed
gvslvol	Living Standard	Standard of living for the old, governments' responsibility
gvslvue	Living Standard	Standard of living for the unemployed, governments' responsibility
gvcldcr	Economics	Child care services for working parents, governments' responsibility
sbstrec	Economics	Social benefits/services place too great strain on economy
sbprvpv	Economics	Social benefits/services prevent widespread poverty
sbeqsoc	Economics	Social benefits/services lead to a more equal society
sbbsntx	Economics	Social benefits/services cost businesses too much in taxes/charges
sblazy	Economics	Social benefits/services make people lazy
sblwcoa	Economics	Social benefits/services make people less willing care for one another
imsclbn	Refugee	When should immigrants obtain rights to social benefits/services
uentrjb	Refugee	Most unemployed people do not really try to find a job
bnlwinc	Refugee	Social benefits only for people with lowest incomes
eduunmp	Security	Spend more on education for unemployed at cost of unemployment benefit
wrkprbf		Benefits for parents to combine work and family even if means higher taxes
basinc		Against or In favour of a basic income scheme
lknemny		How likely not enough money for household necessities next 12 months
vteurmmb	Politics	Would vote for [country] to remain member of European Union or leave
hhmmb		Number of people living regularly as member of household
gndr	Gender	Gender
agea	Age	Age of respondent, calculated
icpart1		Interviewer code, lives with husband/wife/partner

domicil		Domicile, respondent's description
eduyrs	Education	Years of full-time education completed
emplrel	Employment	Employment relation
wrkctra	Employment	Employment contract unlimited or limited duration
wkdcorga	Work	Allowed to decide how daily work is organised
iorgact	Work	Allowed to influence policy decisions about activities of organisation
tporgwk	Work	What type of organisation work/worked for
mbtru	Income	Member of trade union or similar organisation
hincsrca	Income	Main source of household income
hinctnta	Income	Household's total net income, all sources
hincfel	Income	Feeling about household's income nowadays
ipcrtiv	Cognition	Important to think new ideas and being creative
imprich	Cognition	Important to be rich, have money and expensive things
ipeqopt	Cognition	Important that people are treated equally and have equal opportunities
ipshabt	Cognition	Important to show abilities and be admired
impsafe	Cognition	Important to live in secure and safe surroundings
impdiff	Cognition	Important to try new and different things in life
ipfrule	Cognition	Important to do what is told and follow rules
ipudrst	Cognition	Important to understand different people
ipmodst	Cognition	Important to be humble and modest, not draw attention
ipgdtim	Cognition	Important to have a good time
impfree	Cognition	Important to make own decisions and be free
iphlppl	Cognition	Important to help people and care for others well-being
ipsuces	Cognition	Important to be successful and that people recognize achievements
ipstrgv	Cognition	Important that government is strong and ensures safety
ipadvnt	Cognition	Important to seek adventures and have an exciting life
ipbhprp	Cognition	Important to behave properly
iprspot	Cognition	Important to get respect from others
iplylfr	Cognition	Important to be loyal to friends and devote to people close
impenv	Cognition	Important to care for nature and environment
imptrad	Cognition	Important to follow traditions and customs
impfun	Cognition	Important to seek fun and things that give pleasure
dweight	Cognition	Design weight
pspwght		Post-stratification weight including design weight
pweight		Population size weight (must be combined with dweight or pspwght)

Table 2: Descriptive statistics of variables

	France			Britain			Italy			Sweden			Russia		
	Obv.	Mean	ST.D	Obv.	Mean	ST.D	Obv.	Mean	ST.D	Obv.	Mean	ST.D	Obv.	Mean	ST.D
netusoft	951	4.13	1.45	773	4.29	1.27	304	4.19	1.34	686	4.63	1.00	168	3.78	1.65
ppltrst	951	4.66	2.08	773	5.49	2.14	304	5.01	2.15	686	6.42	1.96	168	4.23	2.37
pplfair	951	6.00	1.94	773	5.85	1.98	304	4.89	2.10	686	6.81	1.74	168	5.14	2.30
pplhlp	951	4.81	2.08	773	5.79	1.86	304	4.20	2.05	686	6.17	1.80	168	4.76	2.28
polintr	951	2.49	0.95	773	2.16	0.85	304	2.69	0.85	686	2.15	0.74	168	2.25	0.79
psppsgva	951	2.13	0.89	773	2.54	0.82	304	1.74	0.71	686	2.58	0.83	168	2.07	0.91
actrolga	951	2.05	1.04	773	2.44	1.12	304	2.16	0.98	686	2.63	1.02	168	1.90	0.96
psppipla	951	2.02	0.85	773	2.44	0.80	304	1.78	0.69	686	2.73	0.86	168	1.87	0.82
cptppola	951	2.27	0.90	773	2.47	1.09	304	2.37	0.86	686	2.62	1.00	168	1.97	1.06
trstprl	951	4.17	2.22	773	4.78	2.33	304	3.61	2.43	686	6.23	2.15	168	3.99	2.62
trstlgl	951	5.00	2.31	773	6.01	2.22	304	4.73	2.44	686	6.43	2.02	168	4.05	2.77
trstplc	951	6.47	2.06	773	6.66	2.09	304	6.35	1.88	686	6.81	1.94	168	4.18	2.82
trstplt	951	2.91	2.00	773	3.80	2.20	304	2.47	2.17	686	4.88	1.90	168	3.48	2.54
trstprt	951	2.84	1.96	773	3.93	2.06	304	2.52	2.22	686	4.96	1.86	168	3.60	2.58
trstep	951	3.87	2.23	773	3.70	2.36	304	4.45	2.34	686	4.87	1.99	168	2.88	2.56
trstun	951	4.95	2.35	773	5.59	2.23	304	5.02	2.34	686	6.31	1.93	168	2.95	2.56
lrscale	951	4.94	2.40	773	4.88	1.88	304	4.92	2.25	686	5.11	2.28	168	5.15	1.98
stflife	951	6.52	2.19	773	7.37	1.91	304	7.25	1.77	686	8.07	1.53	168	5.65	2.27
stfeco	951	3.48	1.78	773	5.00	2.01	304	3.77	2.04	686	6.12	1.91	168	3.91	2.17
stfgov	951	3.16	1.98	773	4.66	2.19	304	3.37	2.16	686	4.84	1.96	168	4.63	2.31
stfdem	951	4.27	2.39	773	5.32	2.35	304	4.08	2.25	686	6.45	2.09	168	4.38	2.24
stfedu	951	4.69	2.01	773	5.54	2.00	304	5.21	2.15	686	5.15	1.91	168	4.67	2.03
stfhlth	951	6.19	1.94	773	5.52	2.21	304	5.56	2.16	686	5.63	2.01	168	3.89	2.02
gincdif	951	2.03	1.09	773	2.29	1.00	304	1.90	0.81	686	2.29	0.92	168	2.04	0.87
mnrgtjb	951	4.53	0.93	773	4.39	0.82	304	4.03	1.05	686	4.65	0.68	168	3.40	1.20
freehms	951	1.44	0.85	773	1.63	0.77	304	2.13	0.99	686	1.40	0.70	168	3.63	1.29
hmsfmlsh	951	4.57	0.95	773	4.44	0.82	304	4.00	1.00	686	4.58	0.79	168	2.35	1.27
hmsacld	951	2.66	1.48	773	2.29	1.20	304	3.45	1.23	686	1.91	1.01	168	4.18	1.10
euftf	951	4.96	2.56	773	4.22	2.58	304	5.35	2.97	686	5.04	2.13	168	3.96	2.47
imsmetn	951	2.02	0.76	773	2.09	0.72	304	2.21	0.83	686	1.61	0.59	168	2.26	0.92
imdfetn	951	2.29	0.84	773	2.23	0.77	304	2.53	0.87	686	1.69	0.64	168	2.70	0.89
impcntr	951	2.31	0.87	773	2.30	0.82	304	2.51	0.86	686	1.77	0.68	168	2.90	0.85
imbgeco	951	4.83	2.46	773	5.84	2.29	304	4.74	2.57	686	5.94	2.26	168	4.21	2.40
imueclt	951	5.35	2.73	773	5.84	2.56	304	4.76	2.67	686	7.08	2.28	168	3.77	2.60
imwbcnt	951	4.90	2.16	773	5.51	2.34	304	3.74	2.24	686	6.45	2.18	168	3.42	2.25
happy	951	7.26	1.69	773	7.69	1.71	304	7.42	1.61	686	7.97	1.48	168	6.21	1.95
sclmeet	951	5.29	1.38	773	4.82	1.50	304	4.89	1.33	686	5.49	1.36	168	4.26	1.62
inprdsc	951	2.70	1.38	773	3.22	1.36	304	2.18	1.14	686	3.75	1.26	168	2.26	1.29
sclact	951	2.92	0.83	773	2.72	0.93	304	2.97	0.63	686	2.87	0.85	168	2.76	1.07
crmvct	951	1.77	0.42	773	1.84	0.36	304	1.83	0.37	686	1.75	0.43	168	1.91	0.29
aesfdrk	951	1.92	0.87	773	1.89	0.79	304	2.21	0.81	686	1.69	0.71	168	2.05	0.74
health	951	2.31	0.89	773	2.02	0.93	304	2.05	0.76	686	1.89	0.82	168	2.73	0.66
hlthhmp	951	2.65	0.62	773	2.65	0.64	304	2.88	0.36	686	2.68	0.56	168	2.70	0.53
atchctr	951	8.11	1.83	773	7.08	2.44	304	8.03	1.94	686	8.02	1.84	168	7.38	2.49
atcherp	951	5.97	2.34	773	4.65	2.55	304	6.05	2.16	686	6.35	2.15	168	2.60	2.42
rlgdgr	951	4.27	3.27	773	3.57	3.07	304	5.30	2.69	686	2.81	2.75	168	4.39	2.55
rlgatnd	951	6.14	1.20	773	5.86	1.47	304	5.11	1.54	686	5.93	1.16	168	5.59	1.21
pray	951	5.73	2.02	773	5.26	2.35	304	4.18	2.41	686	6.02	1.82	168	4.96	2.02
dscrgrp	951	1.88	0.32	773	1.85	0.36	304	1.97	0.18	686	1.91	0.28	168	1.95	0.21
gvrfgap	951	2.58	1.32	773	2.65	1.07	304	3.27	1.12	686	2.56	0.93	168	3.68	1.14
rfgfrpc	951	3.25	1.15	773	3.03	0.96	304	2.56	1.02	686	3.34	0.88	168	2.61	1.05
rfgbfml	951	2.79	1.31	773	2.90	1.06	304	2.67	1.07	686	2.59	0.96	168	2.70	1.11
eneffap	951	8.08	2.02	773	7.32	2.42	304	8.61	1.63	686	7.45	2.18	168	6.54	2.26

rdcenr	951	4.44	1.11	773	4.34	1.11	304	4.39	1.02	686	4.10	1.03	168	3.53	1.39
cflsenr	951	7.47	1.86	773	6.46	2.44	304	6.33	2.03	686	7.48	2.44	168	4.55	2.48
elgcoal	951	4.29	0.84	773	3.86	0.86	304	4.20	0.90	686	4.66	0.59	168	3.15	1.05
elgngas	951	3.08	0.89	773	3.05	0.90	304	2.80	1.10	686	3.55	1.02	168	2.61	1.09
elghydr	951	2.05	0.83	773	1.99	0.83	304	1.82	0.89	686	2.15	0.93	168	2.28	1.00
elgnuc	951	3.55	1.10	773	3.64	1.14	304	4.31	1.07	686	3.70	1.18	168	2.96	1.25
elgsun	951	1.79	0.80	773	2.02	0.90	304	1.39	0.71	686	1.89	0.93	168	2.42	1.12
elgwind	951	2.01	0.97	773	2.16	1.00	304	1.69	0.88	686	2.11	1.04	168	2.62	1.24
elgbio	951	2.34	1.08	773	2.71	1.04	304	2.31	1.19	686	2.50	1.08	168	3.45	1.25
wrpwrct	951	2.29	1.02	773	2.15	0.87	304	2.35	0.85	686	1.68	0.69	168	2.66	1.15
wrenexp	951	3.41	0.92	773	3.29	0.84	304	3.27	0.81	686	2.23	0.78	168	3.06	1.13
wrdpimp	951	3.17	0.95	773	3.35	0.89	304	3.31	0.78	686	2.54	0.83	168	2.18	1.21
wrdpfos	951	3.42	0.94	773	3.17	0.90	304	3.35	0.85	686	2.77	0.90	168	2.67	1.11
wrntdis	951	3.13	1.00	773	2.57	0.92	304	2.89	0.98	686	2.09	0.79	168	2.88	1.08
wrinspw	951	2.93	0.96	773	2.61	0.89	304	2.71	0.88	686	2.06	0.68	168	2.61	1.07
wrtcfl	951	2.83	0.97	773	2.46	0.86	304	2.49	0.82	686	2.08	0.67	168	2.99	1.01
wrtratc	951	3.20	1.24	773	2.71	1.10	304	2.58	1.04	686	2.25	0.90	168	2.83	1.20
clmchng	951	1.34	0.51	773	1.41	0.58	304	1.32	0.49	686	1.43	0.55	168	1.67	0.63
clmthgt2	951	3.54	1.10	773	3.35	0.95	304	3.30	0.88	686	3.27	1.02	168	2.69	0.90
wrclmch	951	3.31	0.90	773	3.01	0.87	304	3.45	0.74	686	2.86	0.82	168	2.77	0.90
ccgdbd	951	2.87	2.23	773	3.52	2.20	304	3.08	2.19	686	2.82	1.84	168	3.82	2.04
lkredcc	951	5.72	2.05	773	5.78	2.15	304	6.19	2.02	686	6.40	2.11	168	4.93	2.28
lklmten	951	4.11	1.83	773	3.71	1.94	304	4.61	2.15	686	4.88	2.08	168	3.95	2.20
gvsrdcc	951	4.22	1.97	773	4.25	2.02	304	4.60	2.04	686	5.05	2.06	168	4.43	2.27
ownrdcc	951	4.79	2.45	773	4.35	2.54	304	5.00	2.32	686	4.52	2.68	168	3.80	2.43
inctxff	951	3.43	1.18	773	3.08	1.19	304	3.38	1.24	686	2.39	1.18	168	3.30	1.11
sbsrnen	951	2.05	0.99	773	2.18	1.04	304	1.92	1.02	686	1.70	0.86	168	2.28	1.05
banhhap	951	2.28	1.15	773	2.49	1.17	304	2.05	0.90	686	2.60	1.20	168	2.79	1.05
dfincac	951	2.95	1.19	773	2.73	1.07	304	3.30	1.11	686	2.96	1.05	168	3.37	0.98
smdfslv	951	2.45	1.07	773	2.62	0.97	304	2.16	0.83	686	2.43	0.87	168	2.38	1.00
uemplwk	951	5.64	2.85	773	5.10	2.89	304	6.93	2.71	686	3.46	1.98	168	5.12	2.85
slvpens	951	4.26	1.95	773	4.98	2.14	304	4.29	1.98	686	4.40	1.99	168	2.99	2.40
slvuemp	951	4.35	1.83	773	4.53	1.87	304	2.74	1.73	686	4.43	1.58	168	2.46	2.22
gvslvol	951	7.86	1.63	773	7.96	1.53	304	8.28	1.43	686	8.10	1.48	168	8.38	2.02
gvslvue	951	6.22	2.09	773	5.94	2.07	304	7.33	1.91	686	7.05	1.80	168	6.13	2.63
gvcldcr	951	7.49	1.92	773	7.04	2.15	304	8.11	1.61	686	7.93	1.75	168	7.45	2.38
sbstrec	951	2.73	1.19	773	2.64	1.02	304	3.15	0.98	686	3.15	0.95	168	3.63	0.93
sbprvpv	951	2.39	0.99	773	2.56	0.88	304	2.70	1.03	686	2.35	0.78	168	3.13	1.08
sbeqsoc	951	2.69	1.09	773	3.00	0.94	304	2.56	0.94	686	2.40	0.84	168	3.51	1.00
sbbsntx	951	2.69	1.19	773	3.14	0.99	304	2.91	1.04	686	3.23	0.94	168	3.63	0.91
sblazy	951	2.71	1.22	773	2.60	1.07	304	3.19	1.04	686	3.12	1.05	168	3.63	1.07
sblwcoa	951	2.71	1.18	773	2.89	1.02	304	3.20	1.04	686	3.23	0.93	168	3.65	1.04
imsclbn	951	2.99	1.01	773	3.11	0.84	304	3.23	1.08	686	2.72	1.11	168	3.48	0.85
uentrjb	951	2.96	1.18	773	3.11	1.04	304	2.80	1.04	686	3.59	0.87	168	3.05	1.08
bnlwinc	951	2.16	0.82	773	2.44	0.75	304	2.50	0.82	686	2.20	0.91	168	2.67	0.68
eduunmp	951	2.75	0.74	773	2.90	0.65	304	2.93	0.62	686	2.70	0.81	168	2.50	0.70
wrkprbf	951	2.47	0.70	773	2.66	0.68	304	2.61	0.65	686	2.44	0.82	168	2.67	0.72
basinc	951	2.41	0.78	773	2.47	0.77	304	2.46	0.75	686	2.10	0.94	168	2.76	0.69
lknemny	951	2.11	0.85	773	1.85	0.84	304	2.13	0.86	686	1.36	0.63	168	2.39	0.95
vteurmmb	951	2.26	5.78	773	1.59	1.68	304	1.58	3.17	686	2.41	6.14	168	3.70	7.97
hhmmb	951	2.23	1.18	773	2.27	1.17	304	2.82	1.21	686	2.50	1.24	168	2.49	1.21
gndr	951	1.52	0.50	773	1.52	0.50	304	1.42	0.49	686	1.45	0.50	168	1.52	0.50
agea	951	50.41	16.82	773	50.98	17.43	304	48.41	15.88	686	50.21	17.28	168	44.78	16.93
icpart1	951	1.44	0.50	773	1.44	0.50	304	1.38	0.49	686	1.30	0.46	168	1.46	0.50
domicil	951	3.00	1.10	773	2.99	0.95	304	3.32	0.89	686	2.83	1.15	168	2.24	1.26
eduyrs	951	12.95	3.53	773	14.44	3.63	304	12.98	4.06	686	13.82	3.35	168	13.57	2.48

icomdng	951	1.96	0.20	773	1.87	0.34	304	1.99	0.11	686	1.91	0.28	168	1.82	0.38
mnactic	951	3.00	2.39	773	2.94	2.53	304	2.43	2.29	686	2.58	2.35	168	2.87	2.52
icpdwrk	951	1.45	0.50	773	1.39	0.49	304	1.34	0.48	686	1.35	0.48	168	1.39	0.49
emplrel	951	1.02	0.21	773	1.05	0.30	304	1.03	0.25	686	1.03	0.26	168	1.05	0.31
wrkctra	951	1.29	0.57	773	1.34	0.69	304	1.22	0.50	686	1.14	0.39	168	1.27	0.63
estsz	951	3.01	1.43	773	3.25	1.34	304	2.68	1.41	686	2.89	1.27	168	2.93	1.24
wkdcorga	951	6.82	3.02	773	6.91	2.89	304	5.56	3.31	686	7.61	2.42	168	4.66	3.33
iorgact	951	4.37	3.17	773	4.58	3.12	304	5.10	3.22	686	5.31	2.95	168	3.03	2.96
tporgwk	951	3.42	1.12	773	3.28	1.26	304	3.51	1.09	686	3.01	1.39	168	3.21	0.99
wrkac6m	951	1.96	0.19	773	1.96	0.20	304	1.97	0.18	686	1.93	0.25	168	1.99	0.08
mbtru	951	2.69	0.62	773	2.33	0.78	304	2.65	0.68	686	1.58	0.74	168	2.27	0.71
hincsrca	951	2.40	1.82	773	2.31	1.80	304	1.77	1.40	686	2.02	1.60	168	2.01	1.67
hinctnta	951	5.18	2.68	773	5.58	2.89	304	5.44	2.39	686	6.75	2.82	168	5.83	2.68
hincfel	951	1.89	0.74	773	1.62	0.73	304	1.96	0.80	686	1.34	0.60	168	2.55	0.80
ipcrtiv	951	2.49	1.25	773	2.45	1.22	304	2.49	1.06	686	2.36	1.14	168	2.82	1.37
imprich	951	4.93	1.02	773	4.44	1.16	304	4.00	1.17	686	4.48	1.12	168	3.65	1.43
ipeqopt	951	1.94	1.07	773	2.12	1.06	304	2.11	0.88	686	1.96	0.98	168	2.51	1.13
ipshabt	951	3.66	1.43	773	3.28	1.36	304	2.73	1.23	686	3.60	1.34	168	2.98	1.44
impsafe	951	2.78	1.42	773	2.56	1.30	304	2.26	1.07	686	3.25	1.29	168	2.51	1.23
impdiff	951	3.03	1.43	773	2.98	1.36	304	2.68	1.19	686	3.10	1.29	168	3.13	1.35
ipfrule	951	4.12	1.44	773	3.53	1.43	304	2.75	1.19	686	3.33	1.36	168	3.25	1.33
ipudrst	951	2.32	1.14	773	2.20	0.97	304	2.25	0.90	686	2.12	0.90	168	2.63	1.24
ipmodst	951	2.35	1.17	773	2.61	1.16	304	2.20	0.90	686	2.70	1.12	168	2.85	1.27
ipgdtim	951	2.36	1.18	773	3.46	1.35	304	3.49	1.27	686	2.86	1.18	168	3.13	1.28
impfree	951	2.42	1.33	773	2.12	1.08	304	2.20	0.96	686	2.31	1.13	168	2.43	1.22
iphlppl	951	2.22	1.13	773	1.93	0.87	304	2.18	0.87	686	2.07	0.91	168	2.54	1.12
ipsuces	951	4.25	1.29	773	3.55	1.35	304	2.35	1.02	686	4.05	1.31	168	3.10	1.37
ipstrgv	951	2.45	1.25	773	2.31	1.19	304	2.05	1.02	686	3.11	1.27	168	2.25	1.17
ipadvnt	951	4.28	1.38	773	3.69	1.39	304	3.90	1.36	686	3.82	1.36	168	3.77	1.51
ipbhprp	951	2.63	1.34	773	2.78	1.32	304	2.44	1.00	686	3.21	1.31	168	2.89	1.21
iprspot	951	3.82	1.42	773	3.49	1.29	304	2.65	1.13	686	3.46	1.26	168	2.95	1.37
iplylfr	951	1.85	0.94	773	1.90	0.89	304	2.02	0.76	686	1.79	0.80	168	2.24	1.00
impenv	951	2.29	1.19	773	2.27	1.09	304	1.84	0.84	686	2.29	1.03	168	2.21	0.98
imptrad	951	3.45	1.58	773	3.04	1.48	304	2.32	1.19	686	3.28	1.43	168	2.50	1.24
impfun	951	3.06	1.36	773	2.89	1.21	304	3.06	1.28	686	2.55	1.10	168	3.26	1.46
dweight	951	1.00	0.41	773	1.00	0.51	304	1.00	0.03	686	1.00	0.00	168	1.03	0.44
pspwght	951	0.95	0.62	773	1.00	0.55	304	0.95	0.19	686	0.97	0.74	168	0.90	1.04