Data Analytics Final Project

Eli Sakov

**Introduction**

Data Analytics encourages the synthesis of methods from across disciplines. For my project, I apply what I have learned in Data Analytics to a problem in Economics. While researching my thesis, I found a paper with an interesting dataset. Blanchard and Woflers (2000) construct a dataset of unemployment figures and labor market institutions and policies. The dataset was also used in a paper by Juan F. Jimeno and Diego Rodriguez-Palenzuela (2003), so it is a research-grade dataset. This is an interesting real-world dataset with the difficulties of a high-dimensional, longitudinal and cross-sectional dataset. Part of the challenge is that it is sparse, and frankly, small, at about 220 observations with more than 80 variables. Small, high-dimensional datasets are extremely challenging to analyze because we do not have the luxury of a lot of data for extremely powerful models.

The dataset was used to investigate unemployment, macroeconomic variables, and labor market institutions. The goal is to predict unemployment, and the effects of independent variables on it. The econometric analysis use a simplistic models

**Data**

The dataset was chosen because it contains interesting data and has been collected for and used in published research. That means it is reliable enough to draw meaningful conclusions, and professionally collected by domain experts. The necessary data for the dataset are unemployment statistics and measures of labor market institutions. This dataset contains both. The dataset was published with the a (Blanchard and Wolfers, 2000), and was retrieved from the author’s personal website. (The original website is gone, but it is available as it was in the year 2000 through The Internet Archive). The data is documented in the data appendix, available with the data.

The dataset contains more than eighty variables, but the relevant ones to this project are.

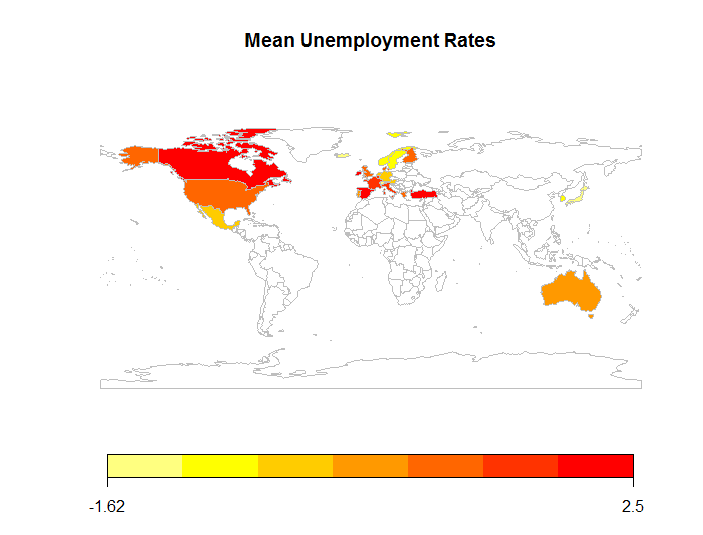
|  |  |
| --- | --- |
| Variable Name | Definition |
| RR | Job replacement Rate |
| Ben | The number of years unemployment benefits last |
| ALMP | A measure of active labor market policy |
| EP | Employment Protection |
| Tax | A measure of the tax wedge |
| Cov | Union contract coverage |
| Den | Union density |
| Coor | Union and employer coordination |
| TFP growth | Total Factor Productivity growth |
| Real rate | The real interest rate |
| LD shift | Labor Demand Shift |
|  |  |
|  |  |
|  |  |
| Dpi | average annual change in the real interest rate |
| e5unr | Average unemployment rate of 5 primary EU countries |
| 15unr | Average unemployment rate of 15 main EU countries |
| Country dummy variables | Dummy variables indicating the 26 countries |
| Time Variables | Dummy variables indicating the years of measure meant: 1965 - 1995 every 5 years. |
| Cty | Country code |

There are many more variables, in the dataset, but these are the important ones. See the appendix for summary statistics for all of them.

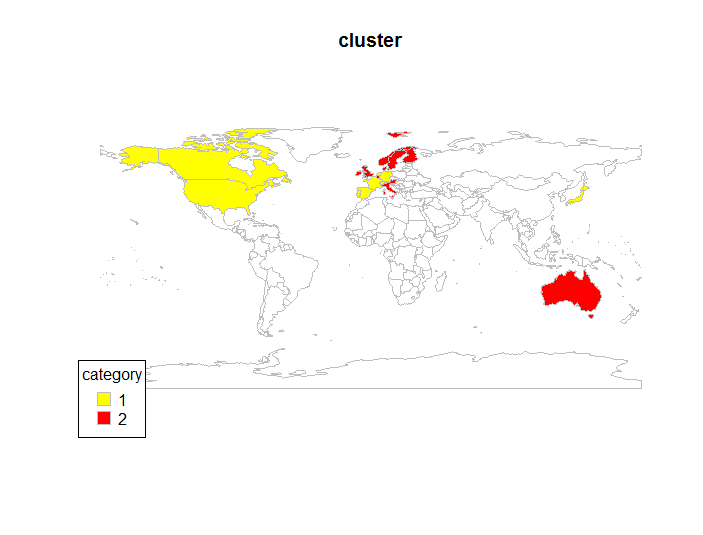
Blanchard

**Exploration**

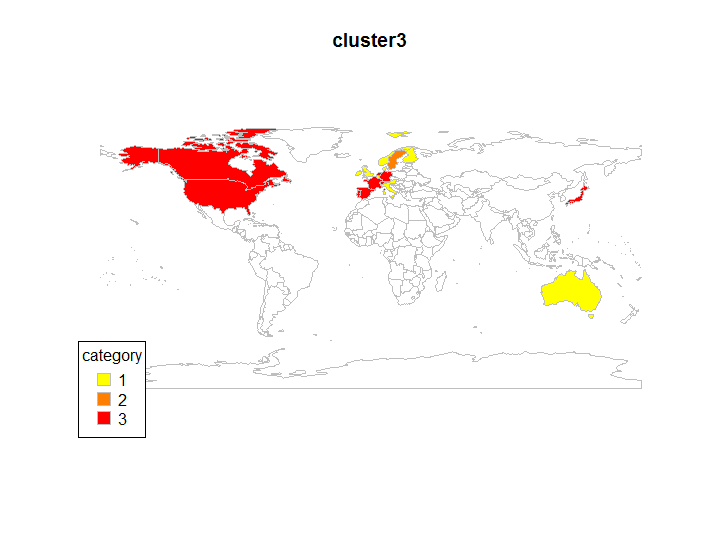
From the original dataset, variables of interest were extracted into a new data frame. They are centered and scaled by subtracting their mean and dividing by their standard deviation. Dummy variables are created for the country and time of each observation.

A good place to start is to display the important variables on their geographies. This graphic shows the mean (scaled) unemployment rate across the period of the dataset. We can see that there are local patterns in long-term unemployment.

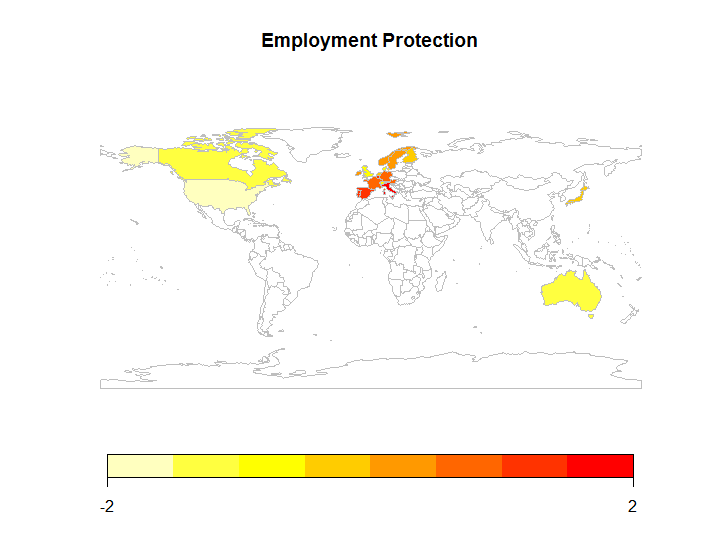
There are too many variables to do this for all of them, but we can cluster them to show similar countries. Beginning with k=2 kmeans clustering shows us the most general division of countries. In the diagrams of cluster, it is important to bear in mind that the coloring does denote intensity.



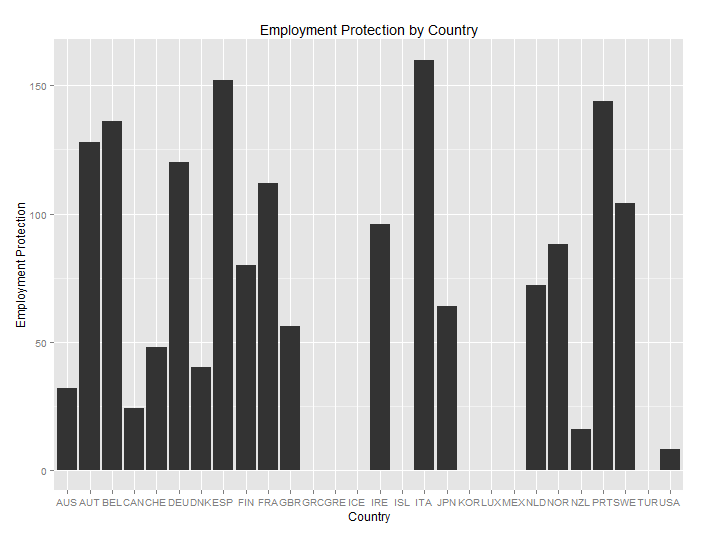
Moving on to 3 clusters, we can see that Sweden is an outlier concerning labor market variables.

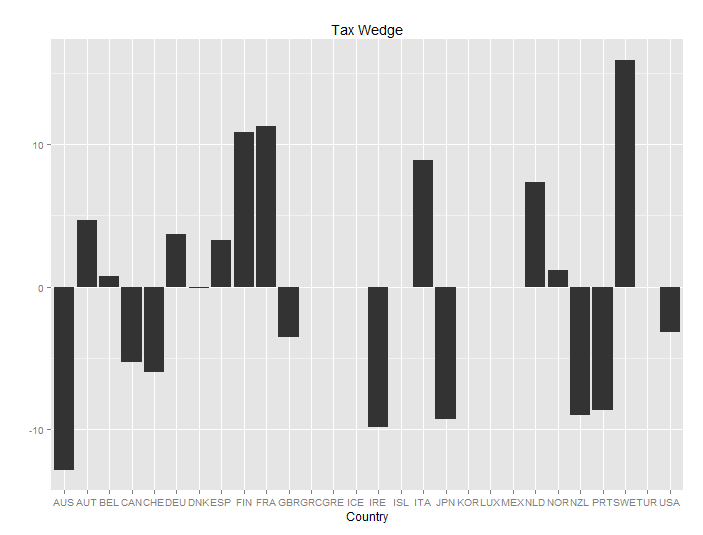


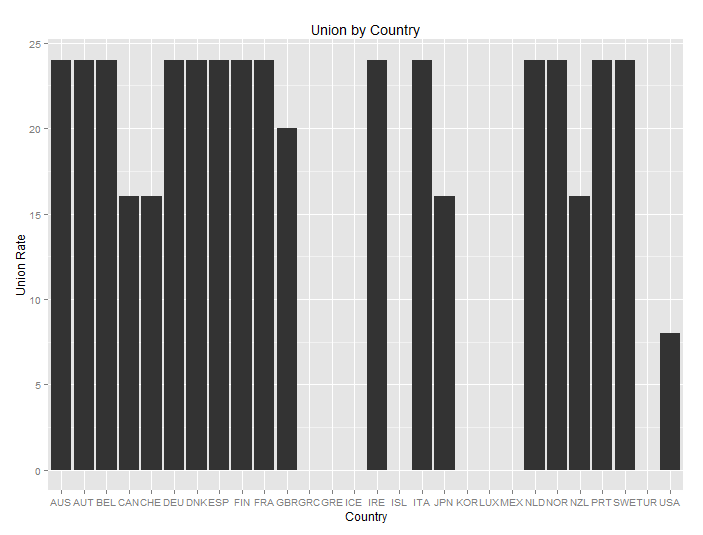
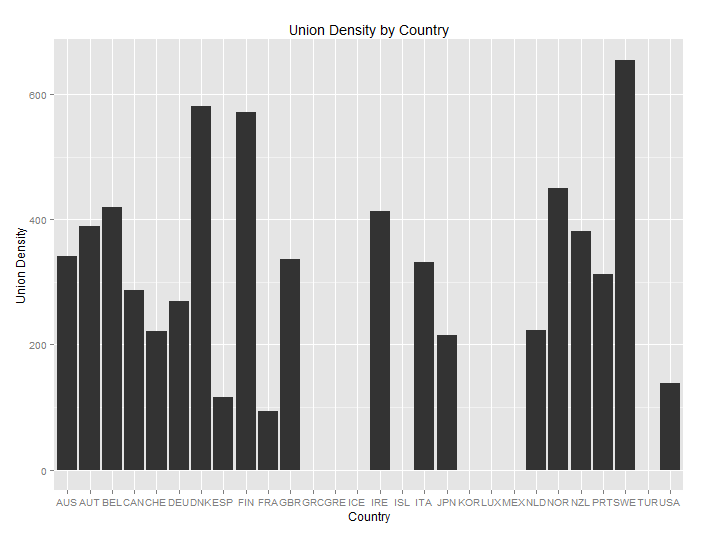
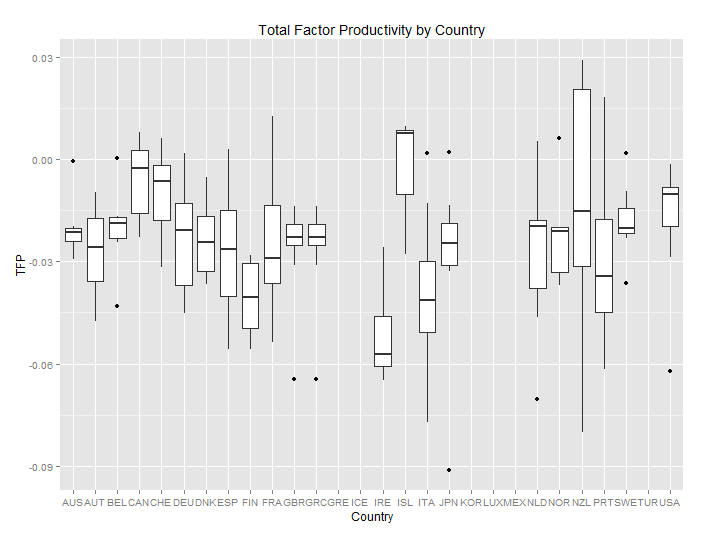
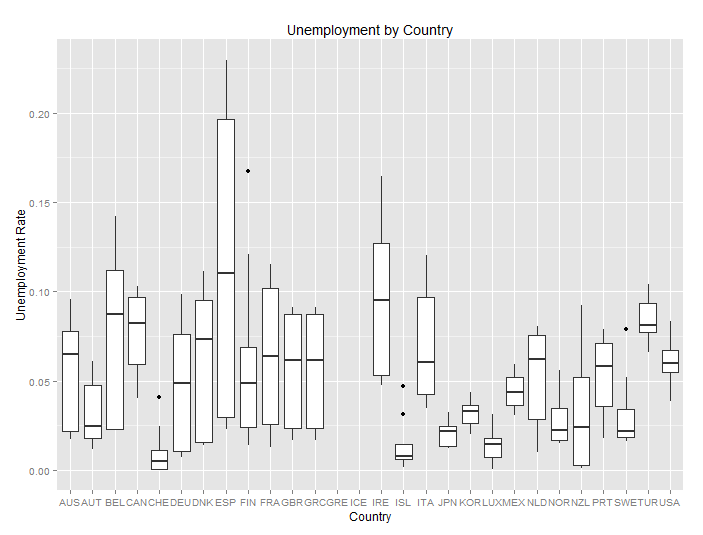
An interesting variable to gain geographic insight from is employment protection. Which seems like a phenomenon in Continental Europe.



Country box plots provide useful information about how the variables vary across each country. Here is what employment protection looks like in a bar chart.







The US is a deviant when it comes to union coverage.

**Models**

Blanchard and Wolfers used a simple, nonlinear least-squares model, but since this is not an economics class, we can experiment with more powerful and less conventional econometric models. Tree-based models that maximize probability can do well with smaller datasets, so I experiment with those in addition to generalized linear mixed models.

To do the most to control for oversitting on a dataset with n <300, models are validated and tuned with leave-one-out cross-validation. That way, the most data can be used for training as possible while getting a generally unbiased error. The forest models don’t need external cross-validation because it is handled internally. Out-of-bag (OOB) error estimates are used for those.

Linear Models are good places to begin investigations, and then move up into more sophisticated models. The first model is a robust GLMM (generalized linear mixed model) absorbing time and country effects.

Linear regression Number of obs = 159

F( 27, 131) = 17.83

Prob > F = 0.0000

R-squared = 0.7725

Root MSE = .02256

------------------------------------------------------------------------------

unr | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

duration | -.0011391 .003492 -0.33 0.745 -.0080471 .005769

empro | -.0026321 .0012313 -2.14 0.034 -.0050678 -.0001964

coord | .0452744 .0128771 3.52 0.001 .0198005 .0707484

uden | .0034005 .001408 2.42 0.017 .0006152 .0061859

union | .0187932 .0076235 2.47 0.015 .0037121 .0338743

almphat | .0026289 .0011579 2.27 0.025 .0003382 .0049196

benefit | -.0100701 .0047832 -2.11 0.037 -.0195325 -.0006077

rrate | .0027632 .0011202 2.47 0.015 .0005471 .0049794

|

Country |

AUT | .0900786 .0291729 3.09 0.002 .0323676 .1477896

BEL | .0207138 .0189355 1.09 0.276 -.0167452 .0581728

CAN | -.0690298 .0326601 -2.11 0.036 -.1336393 -.0044204

CHE | -.0191253 .0102412 -1.87 0.064 -.0393848 .0011342

DEU | .1036409 .0325904 3.18 0.002 .0391694 .1681124

DNK | -.0926074 .0539631 -1.72 0.089 -.1993593 .0141445

ESP | .0939188 .028243 3.33 0.001 .0380474 .1497901

FIN | -.0386368 .0350641 -1.10 0.273 -.108002 .0307283

FRA | .1372319 .0465495 2.95 0.004 .0451459 .2293179

GBR | -.010448 .0100159 -1.04 0.299 -.0302618 .0093659

IRE | -.010378 .0219315 -0.47 0.637 -.0537637 .0330077

ITA | .1323444 .0500731 2.64 0.009 .0332877 .231401

JPN | 0 (omitted)

NLD | 0 (omitted)

NOR | 0 (omitted)

NZL | 0 (omitted)

PRT | 0 (omitted)

SWE | 0 (omitted)

USA | 0 (omitted)

|

Period |

1965 | .0005389 .0070906 0.08 0.940 -.0134879 .0145658

1970 | .0030613 .0070601 0.43 0.665 -.0109053 .0170278

1975 | .0230174 .0063507 3.62 0.000 .0104543 .0355805

1980 | .0486475 .0064931 7.49 0.000 .0358026 .0614923

1985 | .0554815 .0077737 7.14 0.000 .0401032 .0708597

1990 | .0657105 .0071368 9.21 0.000 .0515921 .0798289

1995 | .0733235 .0088419 8.29 0.000 .0558321 .0908149

|

\_cons | -.0826099 .0447467 -1.85 0.067 -.1711296 .0059097

------------------------------------------------------------------------------

It has an F-statistic of 17.83 (p-value <.0001), meaning the model is highly significant. The R-Squared is high at .7725, meaning the model accounts for 77.25% of the variance. The RMSE is .02256, meaning it is a very good fit. 7 country variables were omitted from the regression due to multicollinearity. This is a good model.

The GLMM is also easy to interpret. Duration of benefits has no statistically significant effect, but employment protection, coordination, union density, active labor market policy, and the real interest rate are all significant and all are positively associated with unemployment except benefit and employment protection.

Let’s compare the linear model to a tree based model. Conditional inference trees (ctrees) were chosen because they are unbiased (Hothorn). Tuning was used to select the minimum criterion for a split with the lowest RMSE. The minimum criterion selected is .99 and the RMSE of the final model is 1.035457 with an Rsquared of 37.98%. The Rsquared is lower than the GLMM model, suggesting the GLMM is a better fit.

Conditional Inference Tree

222 samples

21 predictor

No pre-processing

Resampling:

Summary of sample sizes: 33, 33, 33, 33, 33, 33, ...

Resampling results across tuning parameters:

mincriterion RMSE Rsquared

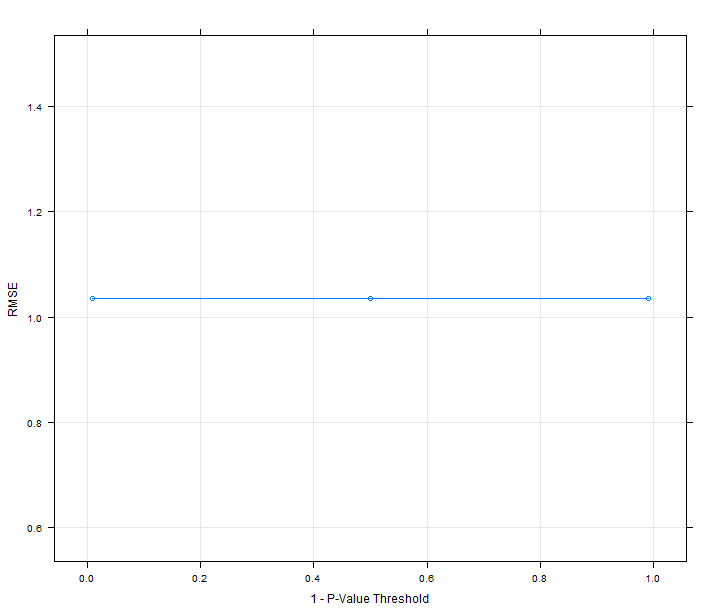
0.01 1.035457 0.3798397

0.50 1.035457 0.3798397

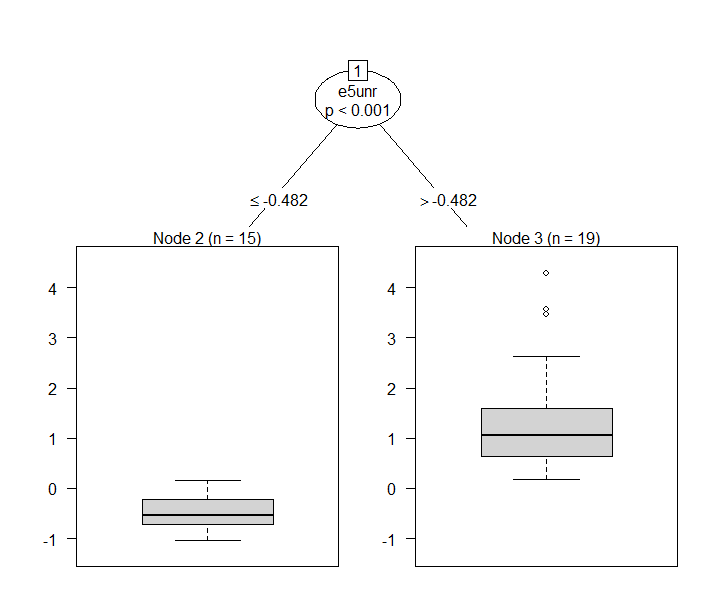
0.99 1.035457 0.3798397

RMSE was used to select the optimal model using the smallest value.

The final value used for the model was mincriterion = 0.99.



The mincriterion parameter appears to have little effect on the error of the model in this model. This is likely because the ctree failed to find more than one place to fit on this data.



This may be due to many weak features. I try using boosting to remedy this.

Boosted Tree

222 samples

21 predictor

No pre-processing

Resampling:

Summary of sample sizes: 33, 33, 33, 33, 33, 33, ...

Resampling results across tuning parameters:

mstop maxdepth RMSE Rsquared

50 1 0.8170098 0.6217724

50 2 0.7898340 0.6373851

50 3 0.7898340 0.6373851

100 1 0.7446554 0.6799717

100 2 0.6721091 0.7331931

100 3 0.6721091 0.7331931

150 1 0.7074300 0.7083018

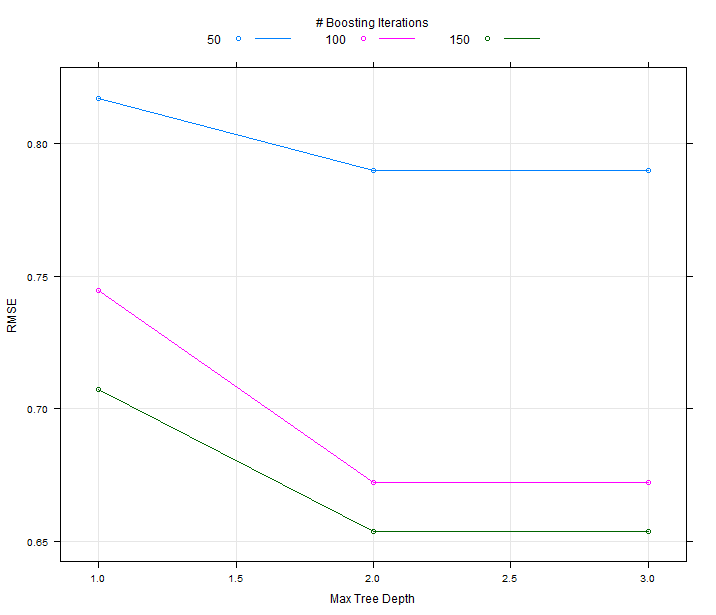
150 2 0.6536884 0.7469631

150 3 0.6536884 0.7469631

Tuning parameter 'nu' was held constant at a value of 0.1

RMSE was used to select the optimal model using the smallest value.

The final values used for the model were mstop = 150, maxdepth = 2 and nu = 0.1.



Using boosting, a tree is able to fit the data much better with an R-Squared of 73.69%. The RMSE should not be used to compare models as it is scale dependant. Tuning picked an mstop of 150, a max depth of 2, and a nu of 0. The library used for boosting does not support displaying the tree. Since more powerful tree models are doing better, next I try a random forest.

Random Forest

222 samples

21 predictor

No pre-processing

Resampling results across tuning parameters:

mtry RMSE Rsquared

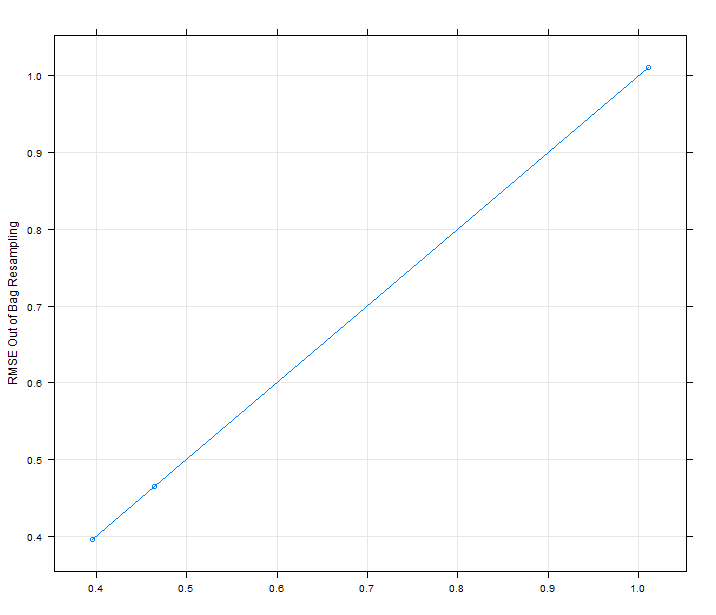
2 1.0101999 0.3951831

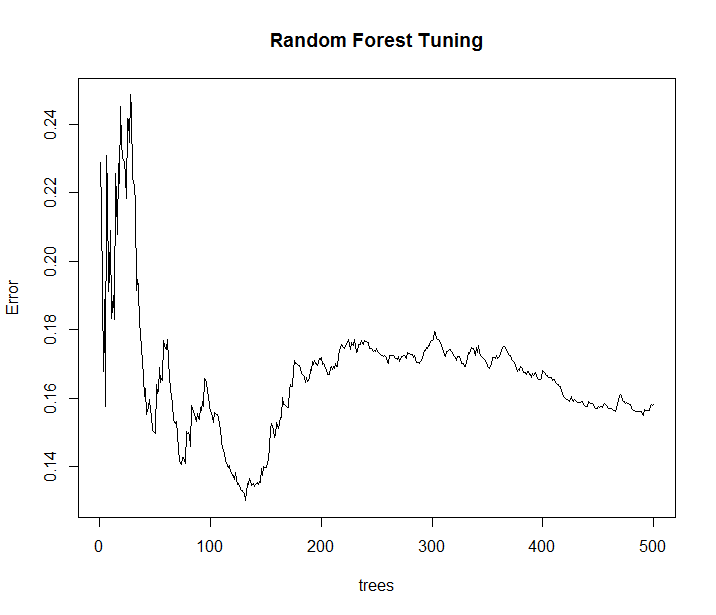
27 0.4650028 0.8718495

52 0.3963508 0.9068959

RMSE was used to select the optimal model using the smallest value.

The final value used for the model was mtry = 52.





The random forest has the best R-Squared of 90.689% and RMSE of 0.39. Tuning selected an mtry (the number of variables selected at each split) of 52. This is a very good model and accounts for 90% of the variance. Random forests can be very robust to overfitting because they implement cross-validation internally using the OOB-error giving an unbiased error.

**Conclusion**

It is ironic that the first linear model is better than most other models. After moving past the GLMM, which trained in less than a second, is easy to interpret, and transparent, I had to work hard to get an equal or better fit. Although the random forest has a 13% better R-Squared, it is only a little better because it is only suitable for prediction whereas the linear model is suitable for inference.

Blanchard and Wolfers try multiple models and fail to get an R-Squared above 87%. While the tuned random forest with boosting does better by that metric, it is not worth the complexity it needed.

**References**

Hothorn, Torsten, Kurt Hornik, and Achim Zeileis. "Unbiased recursive partitioning: A conditional inference framework." *Journal of Computational and Graphical statistics* 15.3 (2006): 651-674.

Jimeno-Serrano, Juan Francisco, and Diego Rodriguez-Palenzuela. "Youth unemployment in the OECD: demographic shifts, labour market institutions, and macroeconomic shocks." (2002).

Blanchard, Olivier, and Justin Wolfers. "The role of shocks and institutions in the rise of European unemployment: the aggregate evidence." The Economic Journal 110, no. 462 (2000): 1-33.

**Appendix**

country cty period ld8 rl tfp

: 14 AUS : 8 Min. :1960 Min. :-0.31839 Min. :-0.19875 Min. :-0.09116

AUSTRALIA: 8 AUT : 8 1st Qu.:1966 1st Qu.:-0.01660 1st Qu.: 0.01036 1st Qu.:-0.03629

AUSTRIA : 8 BEL : 8 Median :1975 Median : 0.00485 Median : 0.03067 Median :-0.02282

BELGIUM : 8 CAN : 8 Mean :1977 Mean : 0.00106 Mean : 0.02712 Mean :-0.02502

CANADA : 8 CHE : 8 3rd Qu.:1985 3rd Qu.: 0.02836 3rd Qu.: 0.04714 3rd Qu.:-0.01365

DENMARK : 8 DEU : 8 Max. :1995 Max. : 0.14551 Max. : 0.18113 Max. : 0.02912

(Other) :168 (Other):174 NA's :68 NA's :59 NA's :68

dpi unr ctyid wadj wstar8

Min. :-0.032942 Min. :0.000038 Min. : 1.0 Min. : 3362 Min. : 3703

1st Qu.:-0.002553 1st Qu.:0.019082 1st Qu.: 7.0 1st Qu.: 17995 1st Qu.: 18429

Median :-0.000040 Median :0.044266 Median :13.5 Median : 40968 Median : 42002

Mean :-0.000095 Mean :0.052399 Mean :13.5 Mean : 759153 Mean : 762179

3rd Qu.: 0.002173 3rd Qu.:0.078779 3rd Qu.:20.0 3rd Qu.: 304502 3rd Qu.: 296084

Max. : 0.028552 Max. :0.229614 Max. :26.0 Max. :10693907 Max. :10780674

NA's :16 NA's :24 NA's :14 NA's :68 NA's :68

nadj gdpb gdpbv ld0 kshare

Min. : 90646 Min. :6.100e+02 Min. :6.615e+03 Min. :-0.42683 Min. :-0.1000

1st Qu.: 3707540 1st Qu.:1.005e+05 1st Qu.:2.328e+05 1st Qu.:-0.05364 1st Qu.: 0.2770

Median : 6944404 Median :3.572e+05 Median :4.589e+05 Median : 0.00127 Median : 0.3090

Mean : 25757820 Mean :4.009e+07 Mean :4.949e+07 Mean :-0.00898 Mean : 0.3089

3rd Qu.: 30885606 3rd Qu.:1.987e+06 3rd Qu.:3.858e+06 3rd Qu.: 0.04205 3rd Qu.: 0.3396

Max. :213573984 Max. :1.382e+09 Max. :1.096e+09 Max. : 0.41812 Max. : 0.6358

NA's :67 NA's :39 NA's :41 NA's :52 NA's :52

gy gn gk irl pi5

Min. :-0.07337 Min. :-0.04946 Min. :0.00025 Min. :0.03232 Min. :0.00653

1st Qu.: 0.01930 1st Qu.:-0.00201 1st Qu.:0.02622 1st Qu.:0.06902 1st Qu.:0.03808

Median : 0.03271 Median : 0.00531 Median :0.03587 Median :0.08791 Median :0.05744

Mean : 0.03339 Mean : 0.00697 Mean :0.03933 Mean :0.10640 Mean :0.09277

3rd Qu.: 0.04465 3rd Qu.: 0.01556 3rd Qu.:0.04720 3rd Qu.:0.11584 3rd Qu.:0.09767

Max. : 0.10953 Max. : 0.05078 Max. :0.11715 Max. :0.72782 Max. :0.76645

NA's :41 NA's :33 NA's :51 NA's :42 NA's :42

pi1 rrate benefit almphat union uden

Min. :-0.002752 Min. :11.00 Min. :0.5000 Min. :-59.265 Min. :1.000 Min. :11.80

1st Qu.: 0.032533 1st Qu.:48.38 1st Qu.:0.9375 1st Qu.:-12.484 1st Qu.:2.375 1st Qu.:27.86

Median : 0.053779 Median :60.00 Median :3.0000 Median : -9.498 Median :3.000 Median :41.70

Mean : 0.088637 Mean :57.45 Mean :2.4487 Mean :-12.020 Mean :2.675 Mean :42.14

3rd Qu.: 0.096784 3rd Qu.:69.25 3rd Qu.:4.0000 3rd Qu.: -7.643 3rd Qu.:3.000 3rd Qu.:51.76

Max. : 0.803024 Max. :90.00 Max. :4.0000 Max. : -2.590 Max. :3.000 Max. :81.80

NA's :16 NA's :62 NA's :62 NA's :62 NA's :62 NA's :62

t coord empro oecdrate duration maxrrate

Min. :29.75 Min. :-6.000 Min. : 1.00 Min. : 0.00 Min. :0.0000 Min. : 0.00

1st Qu.:38.36 1st Qu.:-5.125 1st Qu.: 5.75 1st Qu.:11.97 1st Qu.:0.8715 1st Qu.:30.95

Median :48.12 Median :-4.000 Median :10.50 Median :24.97 Median :2.0800 Median :42.32

Mean :47.64 Mean :-3.950 Mean :10.50 Mean :23.39 Mean :2.2984 Mean :44.36

3rd Qu.:55.05 3rd Qu.:-3.000 3rd Qu.:15.25 3rd Qu.:30.01 3rd Qu.:3.3312 3rd Qu.:60.00

Max. :69.80 Max. :-2.000 Max. :20.00 Max. :67.38 Max. :5.0000 Max. :84.75

NA's :62 NA's :62 NA's :62 NA's :54 NA's :54 NA's :54

maxall rr1 rr2 rr35 rr25 europe

Min. : 0.00 Min. : 0.00 Min. : 0.00 Min. : 0.00 Min. : 0.000 Min. :0.0000

1st Qu.:44.00 1st Qu.:23.52 1st Qu.: 1.65 1st Qu.: 0.00 1st Qu.: 1.462 1st Qu.:0.0000

Median :54.95 Median :34.00 Median :18.60 Median :12.93 Median :15.472 Median :1.0000

Mean :53.79 Mean :37.82 Mean :21.25 Mean :13.62 Mean :15.532 Mean :0.5769

3rd Qu.:65.00 3rd Qu.:55.60 3rd Qu.:32.59 3rd Qu.:23.04 3rd Qu.:24.383 3rd Qu.:1.0000

Max. :92.90 Max. :80.29 Max. :70.00 Max. :60.81 Max. :60.808 Max. :1.0000

NA's :54 NA's :54 NA's :54 NA's :54 NA's :54 NA's :14

eurobig e5unr e15unr avtfp tfpgap

Min. :0.0000 Min. :0.01895 Min. :0.01856 Min. :-0.05161 Min. :-0.06582

1st Qu.:0.0000 1st Qu.:0.02486 1st Qu.:0.02359 1st Qu.:-0.02806 1st Qu.:-0.00982

Median :0.0000 Median :0.06554 Median :0.06224 Median :-0.02389 Median : 0.00149

Mean :0.1923 Mean :0.06542 Mean :0.06136 Mean :-0.02418 Mean : 0.00000

3rd Qu.:0.0000 3rd Qu.:0.10265 3rd Qu.:0.09471 3rd Qu.:-0.01820 3rd Qu.: 0.00972

Max. :1.0000 Max. :0.11934 Max. :0.11193 Max. :-0.00355 Max. : 0.04729

NA's :14 NA's :182 NA's :102 NA's :46 NA's :68

portrev newep Icoun\_2 Icoun\_3 Icoun\_4

Min. :0.00000 Min. :0.000 Min. :0.00000 Min. :0.00000 Min. :0.00000

1st Qu.:0.00000 1st Qu.:1.225 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000

Median :0.00000 Median :2.800 Median :0.00000 Median :0.00000 Median :0.00000

Mean :0.01351 Mean :2.566 Mean :0.03846 Mean :0.03846 Mean :0.03846

3rd Qu.:0.00000 3rd Qu.:4.000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

Max. :1.00000 Max. :4.000 Max. :1.00000 Max. :1.00000 Max. :1.00000

NA's :14 NA's :14 NA's :14

Icoun\_5 Icoun\_6 Icoun\_7 Icoun\_8 Icoun\_9

Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000

1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000

Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000

Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846

3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000

NA's :14 NA's :14 NA's :14 NA's :14 NA's :14

Icoun\_10 Icoun\_11 Icoun\_12 Icoun\_13 Icoun\_14

Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000

1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000

Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000

Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846

3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000

NA's :14 NA's :14 NA's :14 NA's :14 NA's :14

Icoun\_15 Icoun\_16 Icoun\_17 Icoun\_18 Icoun\_19

Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000

1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000

Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000

Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846

3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000

NA's :14 NA's :14 NA's :14 NA's :14 NA's :14

Icoun\_20 Icoun\_21 Icoun\_22 Icoun\_23 Icoun\_24

Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000

1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000

Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000

Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846 Mean :0.03846

3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000

NA's :14 NA's :14 NA's :14 NA's :14 NA's :14

Icoun\_25 Icoun\_26 Iperio\_2 Iperio\_3 Iperio\_4

Min. :0.00000 Min. :0.00000 Min. :0.0000 Min. :0.0000 Min. :0.0000

1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000

Median :0.00000 Median :0.00000 Median :0.0000 Median :0.0000 Median :0.0000

Mean :0.03846 Mean :0.03846 Mean :0.1261 Mean :0.1261 Mean :0.1261

3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000

Max. :1.00000 Max. :1.00000 Max. :1.0000 Max. :1.0000 Max. :1.0000

NA's :14 NA's :14

Iperio\_5 Iperio\_6 Iperio\_7 Iperio\_8 Icoun\_1 Iperio\_1

Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.0000

1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000

Median :0.0000 Median :0.0000 Median :0.0000 Median :0.0000 Median :0.00000 Median :0.0000

Mean :0.1261 Mean :0.1261 Mean :0.1261 Mean :0.1171 Mean :0.03604 Mean :0.1261

3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.0000

Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.0000

zero equnr2 equnr4

Min. :0 Min. :-0.01957 Min. :-0.03981

1st Qu.:0 1st Qu.: 0.02089 1st Qu.: 0.02080

Median :0 Median : 0.04526 Median : 0.04608

Mean :0 Mean : 0.05230 Mean : 0.05219

3rd Qu.:0 3rd Qu.: 0.07773 3rd Qu.: 0.07447

Max. :0 Max. : 0.22209 Max. : 0.21457

NA's :24 NA's :24