Analysis of U.S. Bank Performance During The Pandemic

ANLY 512 Final Project

Team member: Qian Leng, Yihui Liu, Zehui Wu

GEORGETOWN UNIVERSITY

Background

Community banks primarily provide services to the local business in the neighborhood communities which will enhance their personal relationship with the customers. This research project studies the influence that the community banks have on the US economy than the other regular banks and to identify the role that community banks play in recent years.

**Key Words:** *Loan Approval Amount, Programming Analysis, Bank Regulation*

Data Overview

Barplot of Avg. Approval Rate / Lender State

Barplot of Avg. Approval Amount / Lender State (For comparison to the Avg. Approval Rate / Lender State)
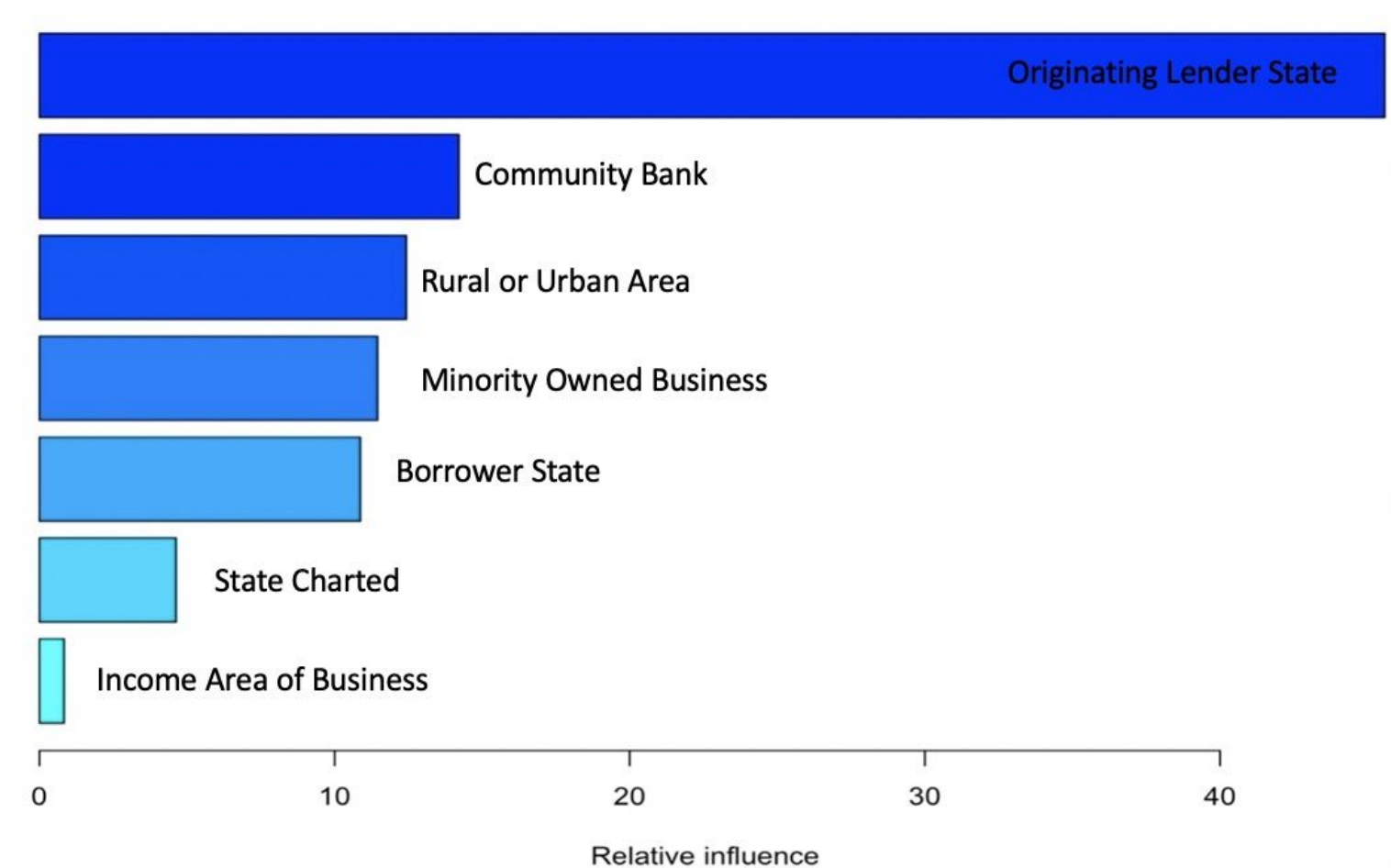
- **Paycheck Protection Program (PPP)** dataset provided by CSBS has been used. The raw dataset has 19 columns and 1048575 rows

- Three new columns which are called "employees ratio", "approval ratio" and "forgiveness ratio" are created to avoid any inaccuracies due to biased sample sizes

Machine Learning – Random Forest

The Random forest is made to compare the accuracy of prediction of decision trees.

Ranking of Variable Importance
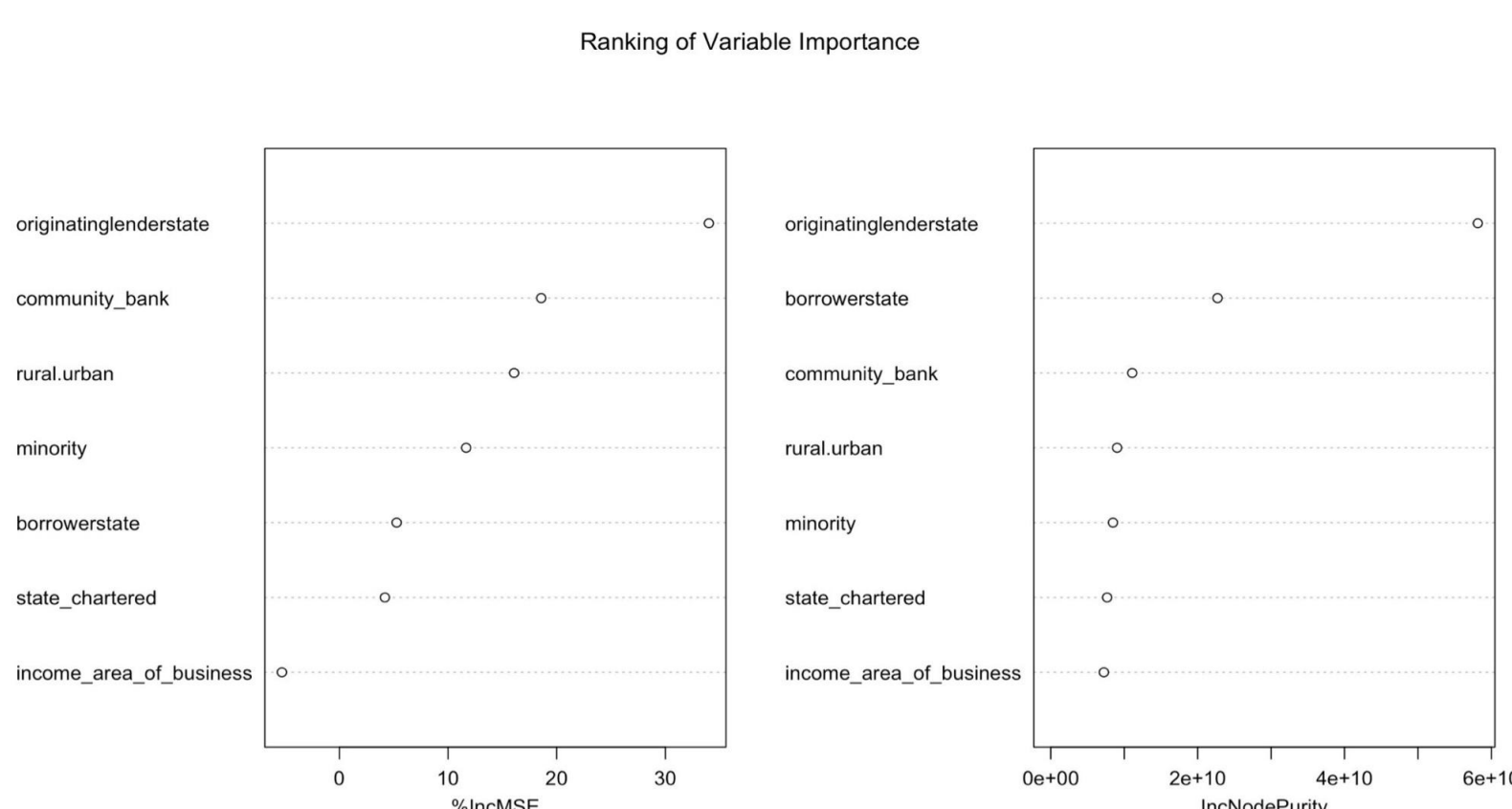
The right relative influence of variables shows directly about how variables influenced the prediction of approval amount. Policy makers can decide to put more emphasis on these variables if they would like to increase the financial level of banks.

Research Highlights

Hypothesis Testing

Originating Lender States	P-value of Chi-Square Test for forgiveness amount and different income region indicator	P-value of Chi-Square Test for forgiveness amount and rural or urban indicator
CA	0.6381022	0.007139012
IL	0.4225591	0.3667032
NC	0.3118421	0.03731832
NJ	0.03944929	0.830098
OH	0.02000392	0.1853785
PA	0.00536135	0.001197444

Regression Analysis

Scatterplot Matrix of Approval/Forgiveness/Jobreported Related Variables
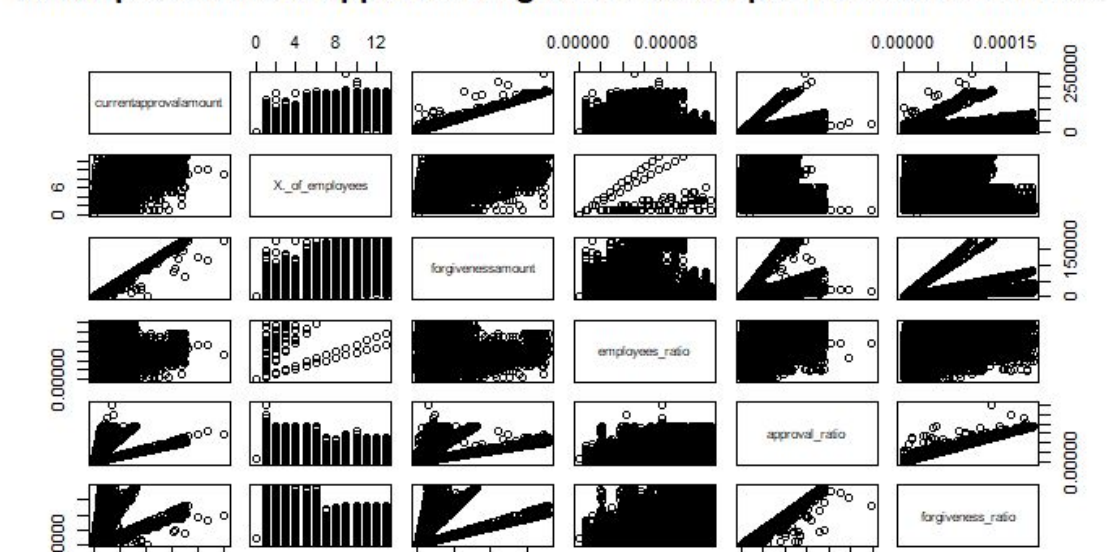
Summary Table for Linear Regression Model		
Model	RMSE	R squared
Model 1	0.0000035	0.9919341
Model 2	0.0000338	0.2359323
Model 3	0.0000035	0.9919049
Model 4	0.0000339	0.2390237
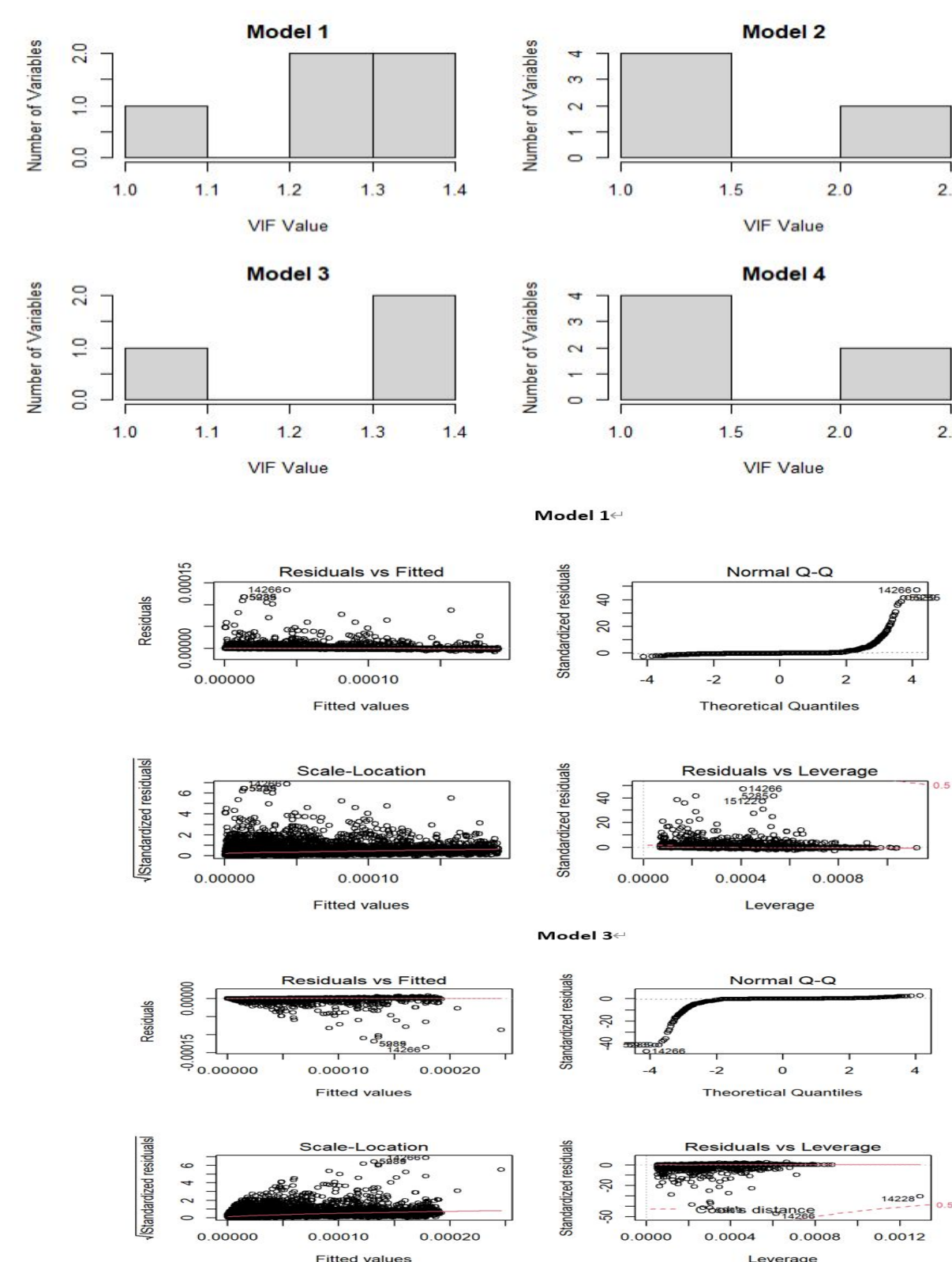
Machine Learning – Boosting Model

Variable	Relative Influence
Originating lender state	45.5814121
Minority	14.2082064
Community_bank	12.4233543
State_chartered	11.4555700
Income_area_of_business	10.8694100
Borrowerstate	4.6265035
Rural.urban	0.8355438

For Boosting model, the variable influence is different than before as the minority owned business is more influential than the community bank. To fairly compare the utility of decision tree, random forest, and boosting model and choose the best model for giving suggestions to policy makers, regulators, and ordinary borrowers, a summary table of RMSE value is created.

Clearly, Random Forest owns the lowest error, and would be the best model to choose.

Machine Learning Methods	RMSE Value
Decision Tree	384945.0923
Random Forest	20707.474665
Boosting Model	59148.19

Results & Next Steps (Neural Network)

By doing **hypothesis testing**, it could be roughly suggested that the PPP loan forgiveness amount of community banks in rural areas is higher than non-community banks, however, this pattern doesn't apply to urban areas.

For **regression analysis**, it could be concluded from the final models that banks with a high forgiveness ratio as well as employees ratio, community banks and urban banks are more likely to have a high approval ratio.
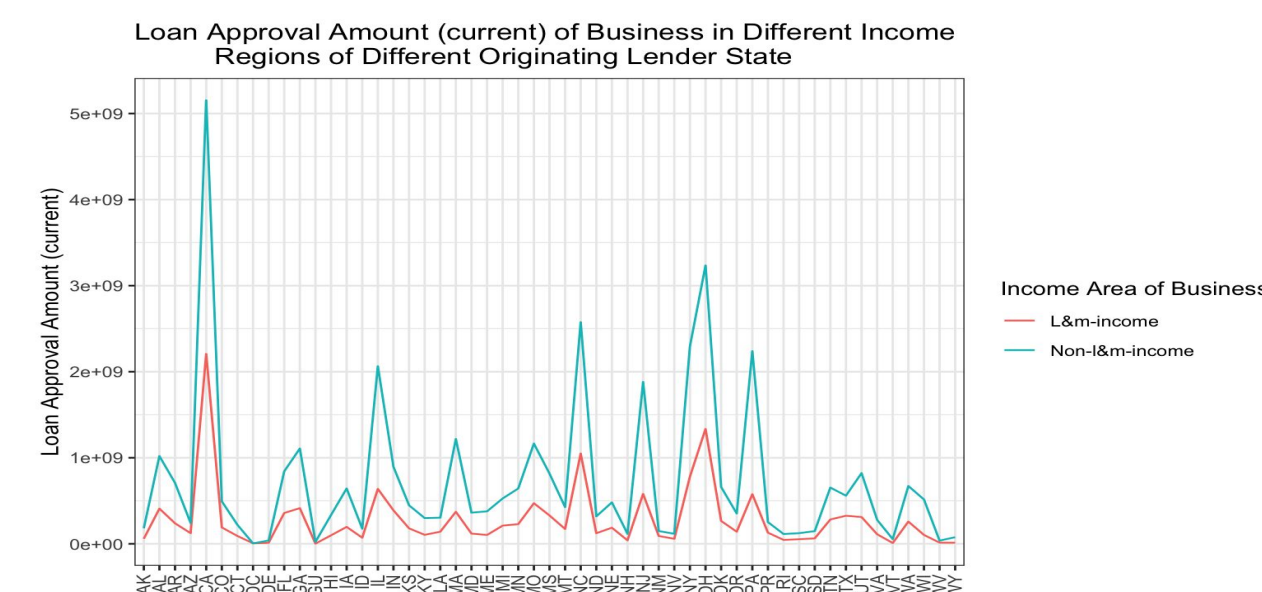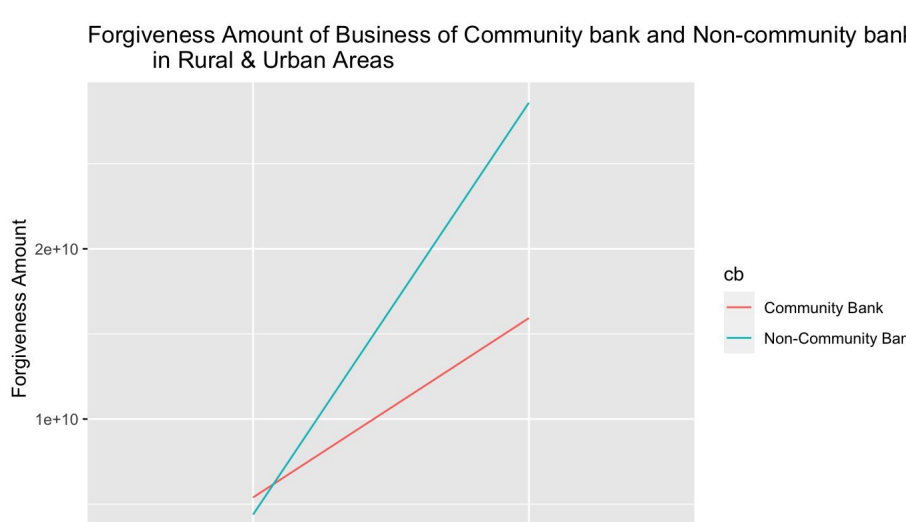
After comparing the RMSE value of **Decision Tree**, **Random Forest**, and **Boosting Modeling**, the prediction of Random Forest was proved to be the most accurate. It can be concluded that the community bank indeed plays a vital role in the financial power of a state.

The **Deep Learning method** can be applied to increase the model prediction on performance of community banks.

Conclusions

I. Policymakers can therefore be suggested by analyzing the advantage of lender states' policy of community banks and setting up more financial support to solve the long-distance borrowing problem

II. Providing more programs on the federal, state, county and city levels which can support and foster minority business enterprises to grow, and can also mitigate the monopoly from large business

III. Making more regulations and offering benefits to community banks to fill the gap between low approval amount and high approval amount and then balancing the distribution (Same for forgiveness amount)

IV. If a government would like to increase the capability of employment ratio, it can be suggested to increase the financial power first. However, this point needs to be argued before enacting other policies