THE UNIVERSITY OF CHICAGO

Application of Machine Learning

Joan Zhang

MScA Capstone Presentation
The Graham School University of Chicago
August 20, 2016, Revised @ Oct, 2017

Agenda

- Introduction
- Methodology
- Findings
- Conclusion

Opportunities of Machine Leaning

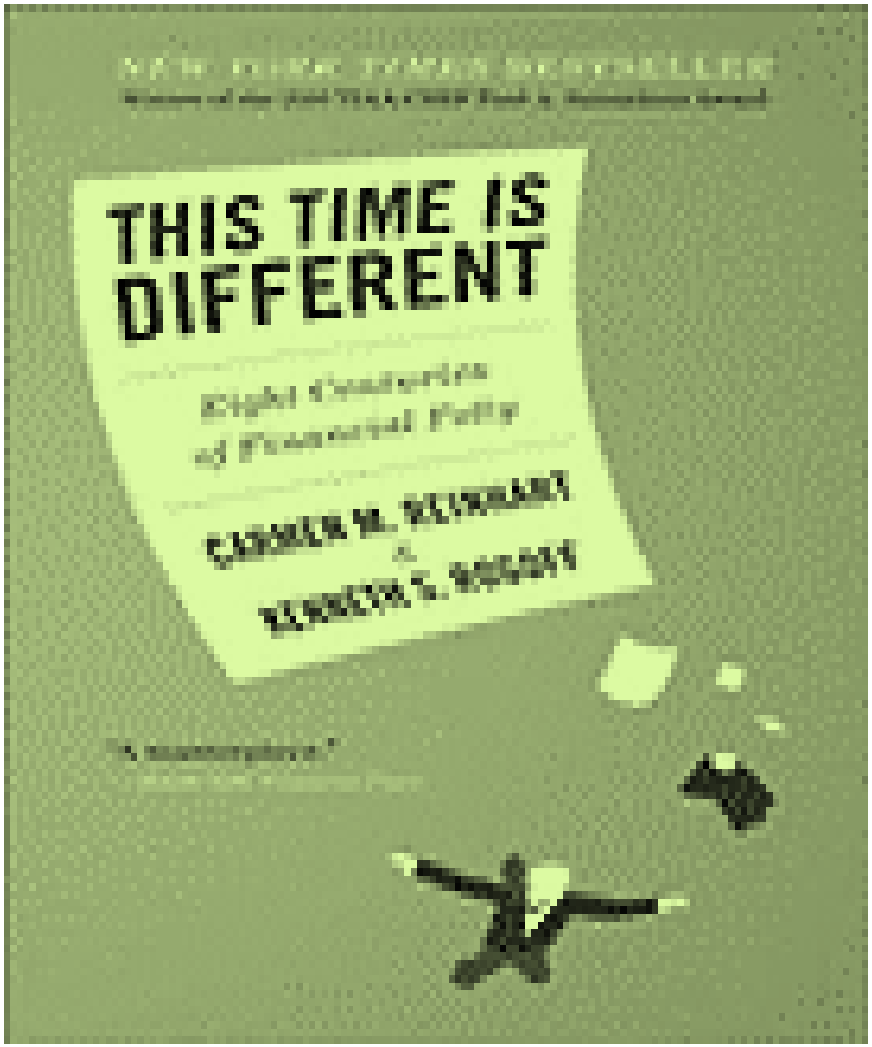
Customer Segmentation

What is Machine Learning (ML)?

- ML is a semi-automated extraction of knowledge from data
- You apply some processes or algorithms to the data using a computer
- ML requires you to make any smart decisions

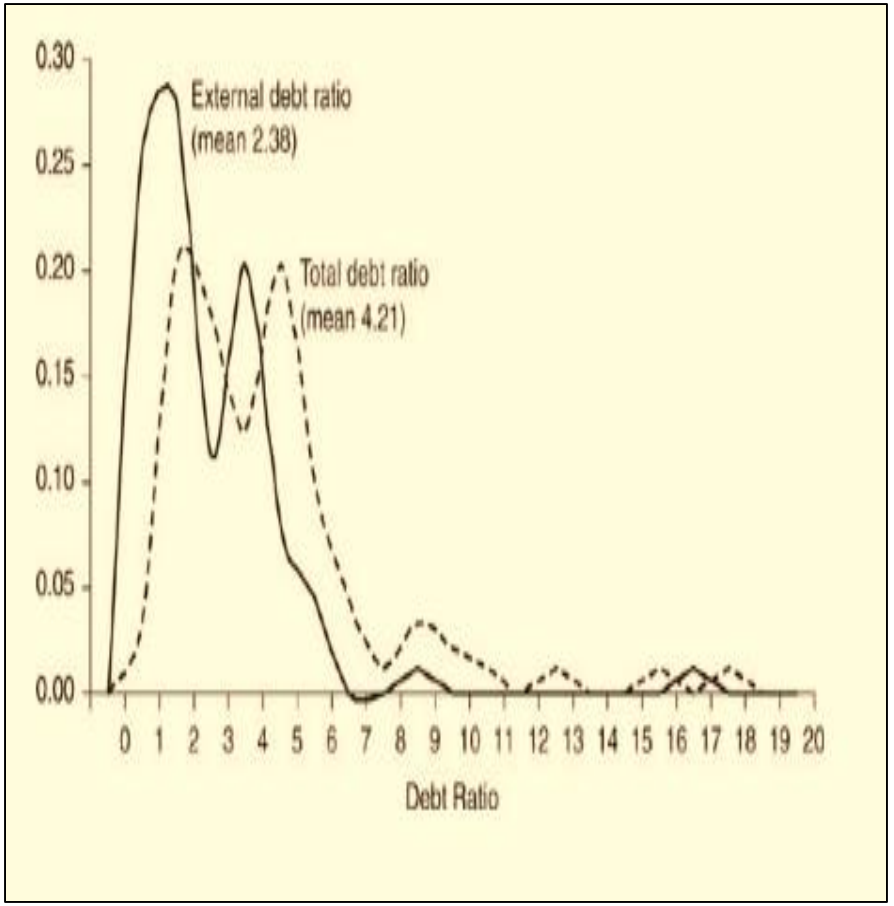
Introduction

- Reoccurrence of financial crises
- Less cautionary steps of towards financial crises
- The "this time is different" syndrome
- More similarities of financial crises



Problem Statement

- Serial default occurs from time to time around the globe
- Domestic debt plays a significant role in default even though external debt is at a very low level



Research Purpose

To examine the consistency of serial defaults throughout history

To examine authors' assertion that domestic debt plays a significant role in default even though external debt is at a very low level

To argue that domestic debt is overlooked

Research Questions

- Is serial default a historical pattern?
- Does domestic debt play a significant role in default?
- What are the correlations among external debt, domestic debt and total debt ?

Variables Identified

- Domestic Debt
- Domestic Debt Ratio
- External Debt
- External Debt Ratio
- Gross National Product (GNP)
- Total Debt Ratio
- Revenue
- Inflation

Methodology

Methodology

Known as predict modeling. For example, predict email as spam or non-spam

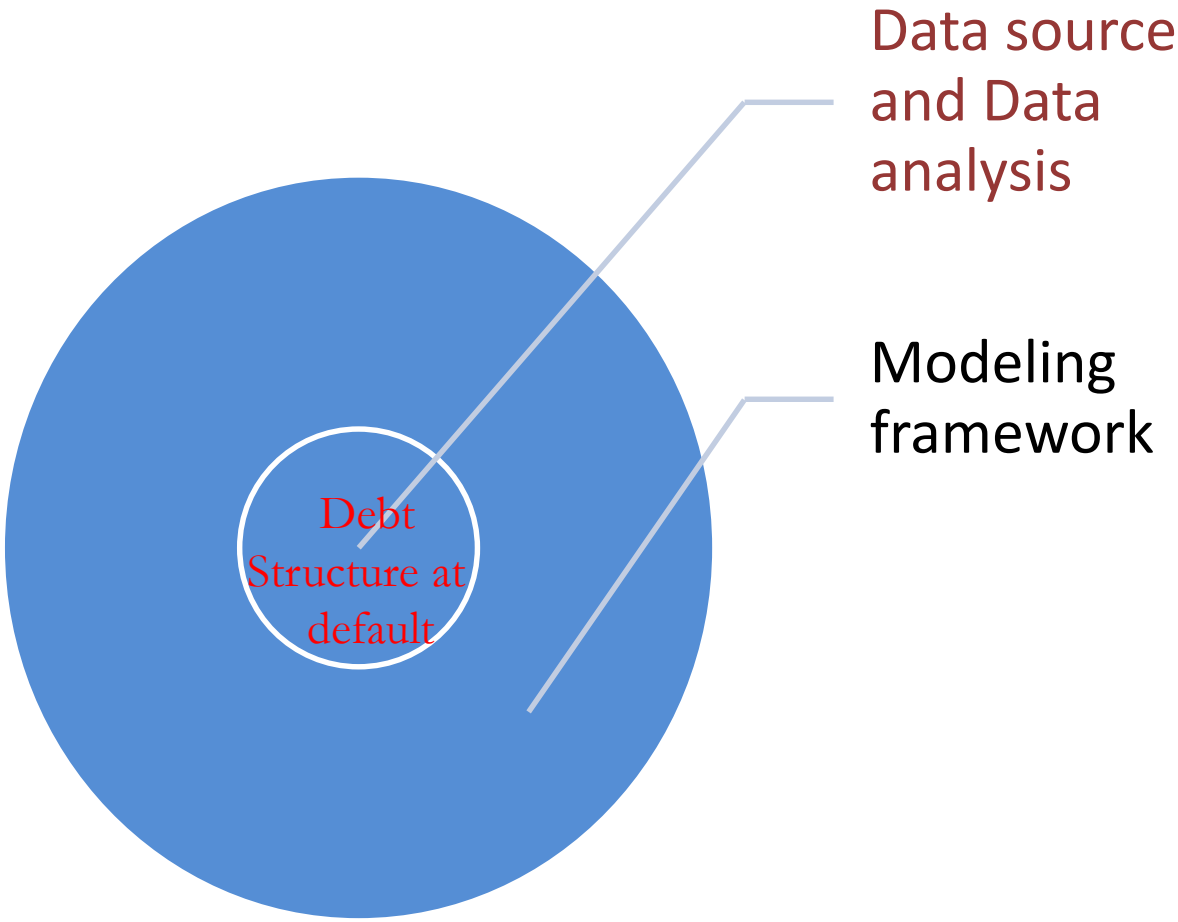
Extracting structure from data. For example, segment customers to prompt productions

Methodology

Analysis of Debt Structure at Default

Analysis of government strategy after default

Analysis of Debt Structure at Default



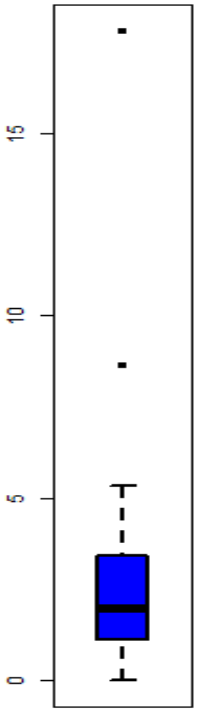
Data Source and Data Analysis

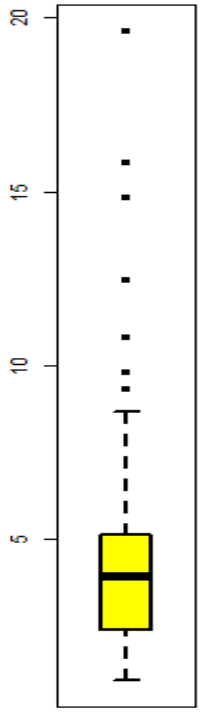
Summary Statistics and Box Plots

Table 1

Summary Statistics

Variable Name	No of Obs	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	StDev
External Debt Ratio	89	0.00	1.12	1.95	2.47	3.44	17.79	2.33
Total Debt Ratio	89	0.94	2.40	3.92	4.37	5.14	19.61	3.26
Domestic Debt Ratio	89	0.09	1.01	1.29	1.90	2.30	10.88	1.81





Pearson Correlation and Spearman Correlation

Table 2

	Pearson	Spearman
External to Domestic Debt Ratio	0.2196669	0.2164136
External to Total Debt Ratio	0.8393286	0.8362263
Domestic to Total Debt Ratio	0.7147191	0.6574123

Histograms

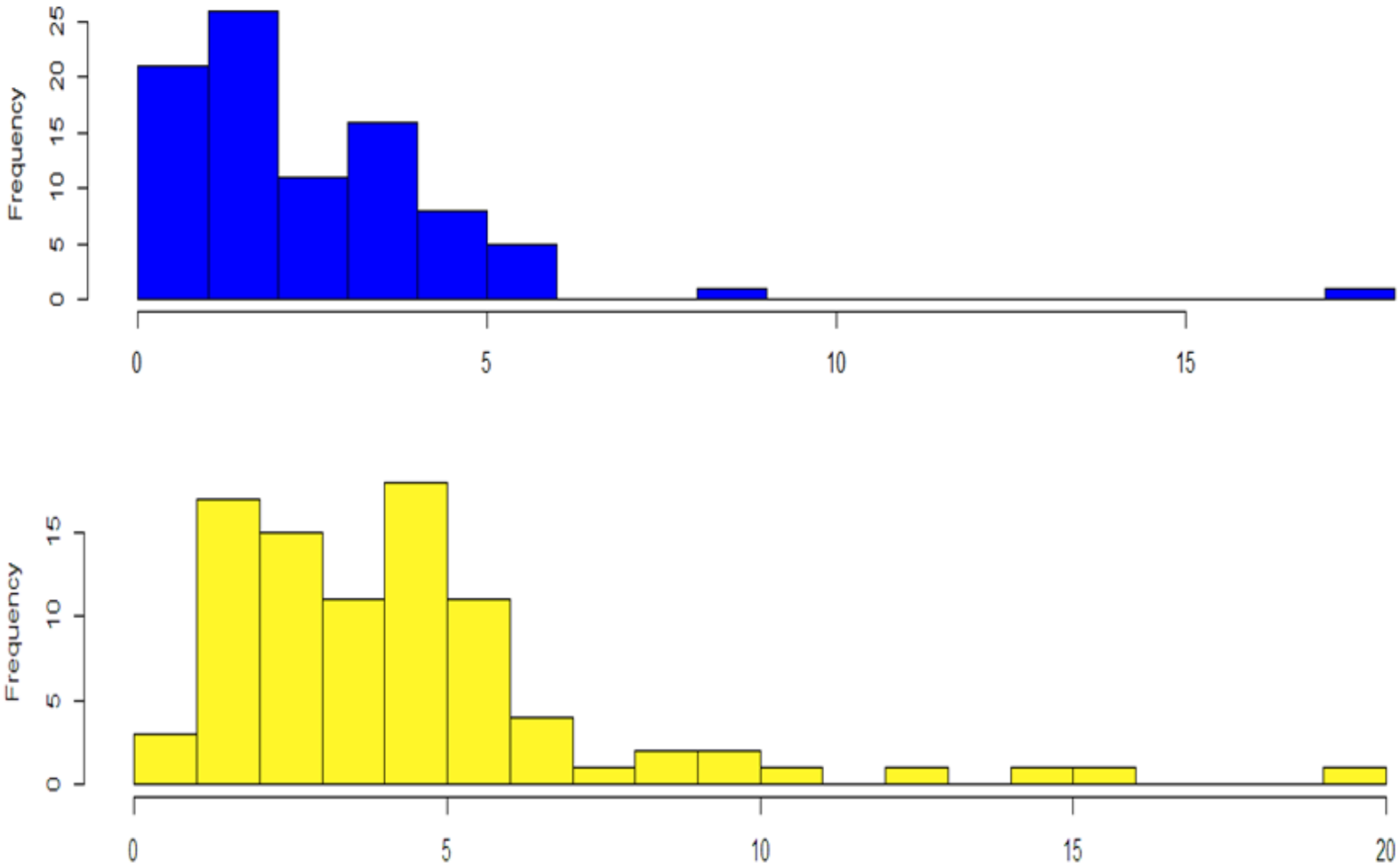
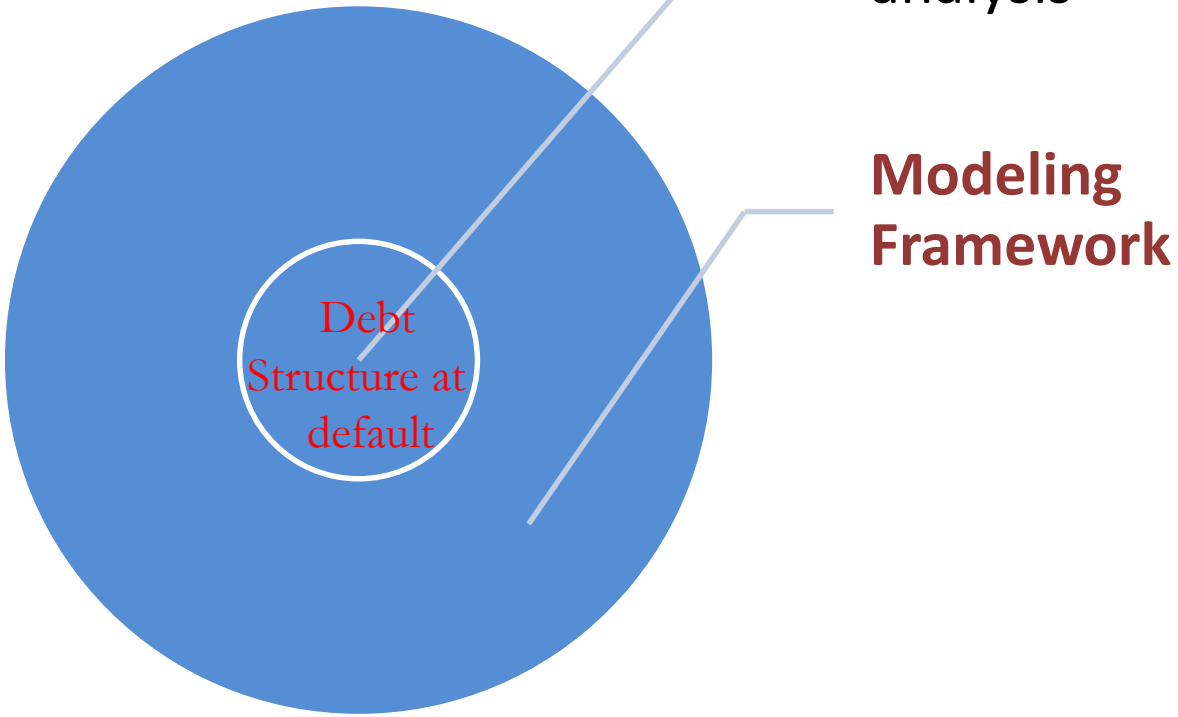


Figure 2. External debt ratio distribution and total debt ratios distribution

Analysis of Debt Structure at Default



Modeling Framework

Data Preparation

- Outliers in dataset
- Log transformation

Assess the Clustering Tendency

Hopkins Statistic

- Null hypothesis: the dataset is uniformly distributed (i.e., no meaningful clusters, H is about 0.5)
- Alternative hypothesis: the dataset is not uniformly distributed (i.e., contains meaningful clusters)

Data Visualization

Hopkins test return H value = 0.23

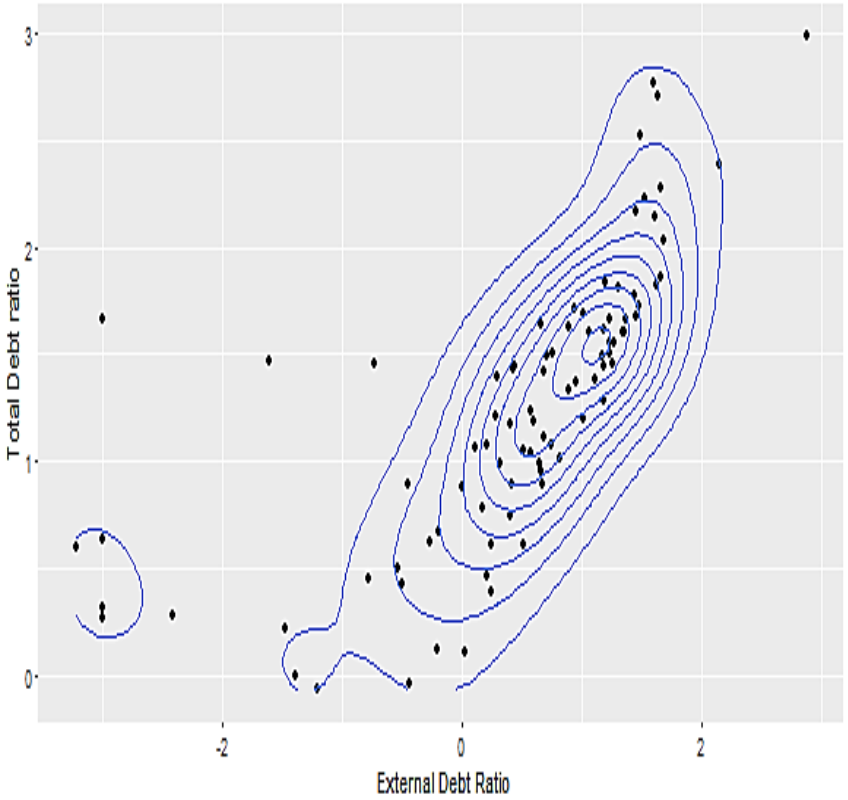


Figure 9. Visualization of external debt ratio and total debt ratio on two dimensions

Random uniformly distributed data

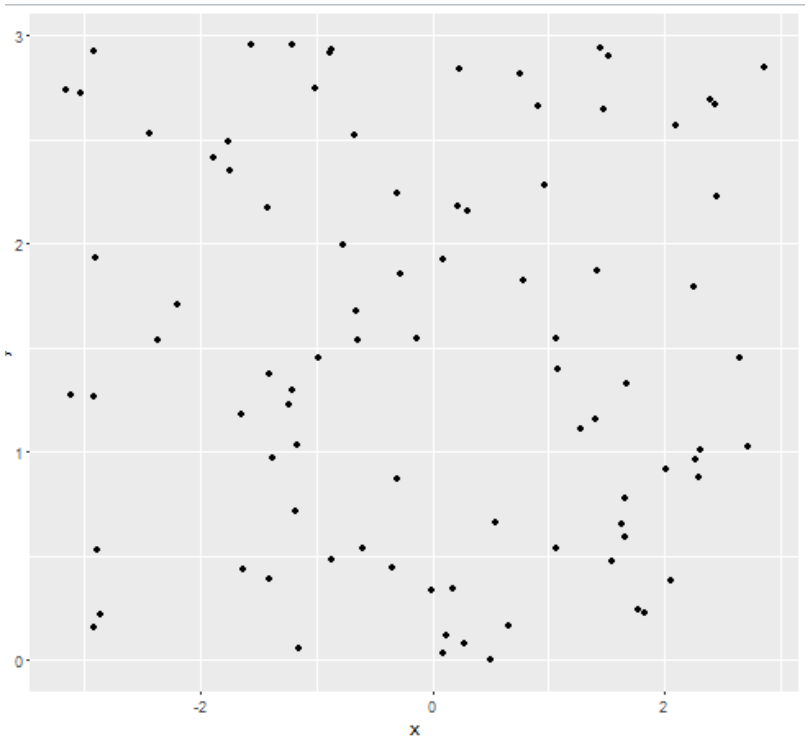


Figure 10. Random Uniformly Distributed Non-Clustering Dataset

Mclust

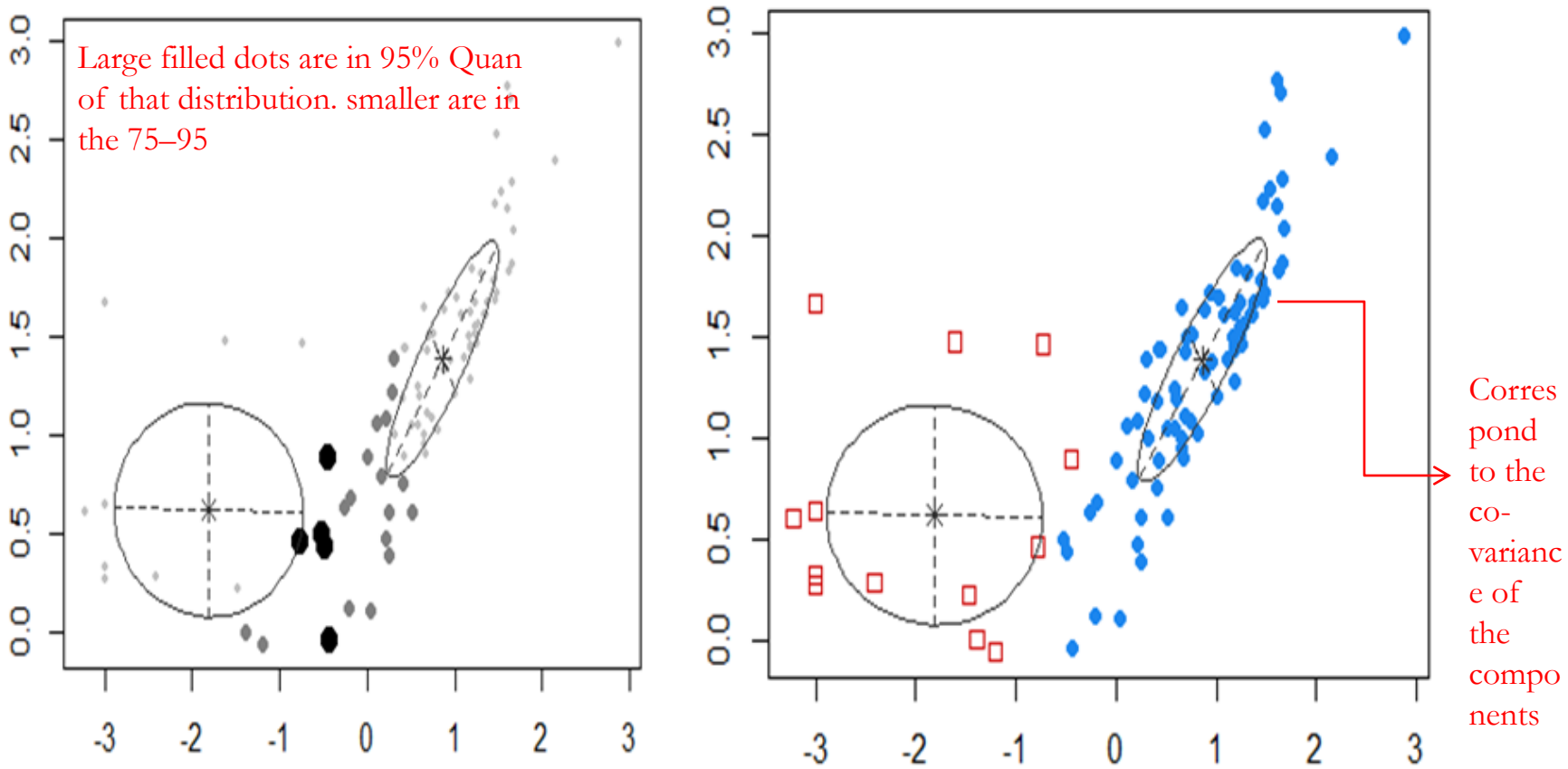
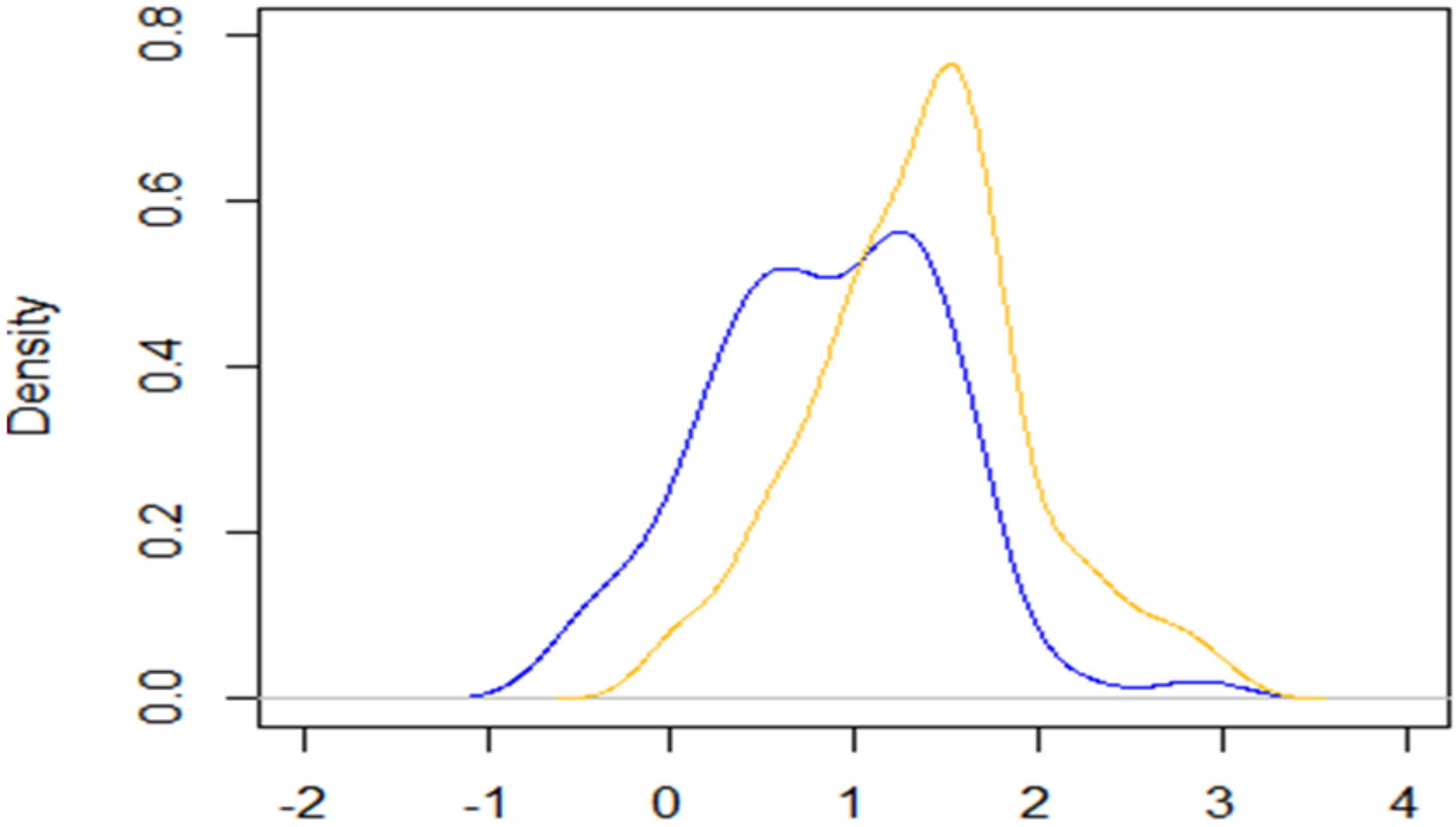


Figure 11. Plots of uncertainty and classification

Analysis of Clustering Output

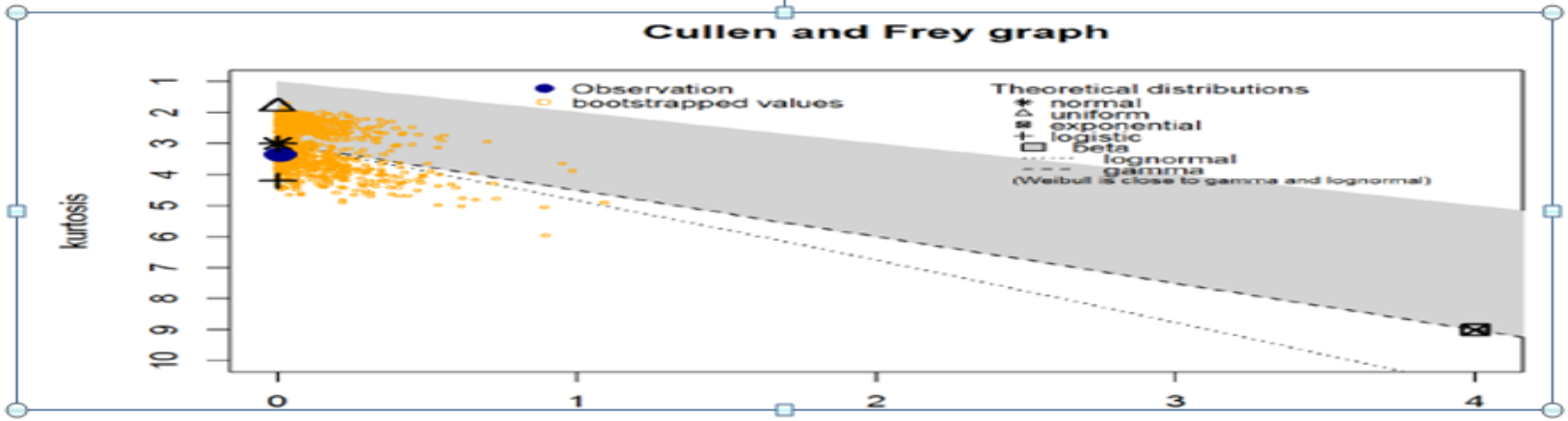
- Distribution from cluster 1 group
- Distribution from cluster 2 group

Density Plot of Cluster 1

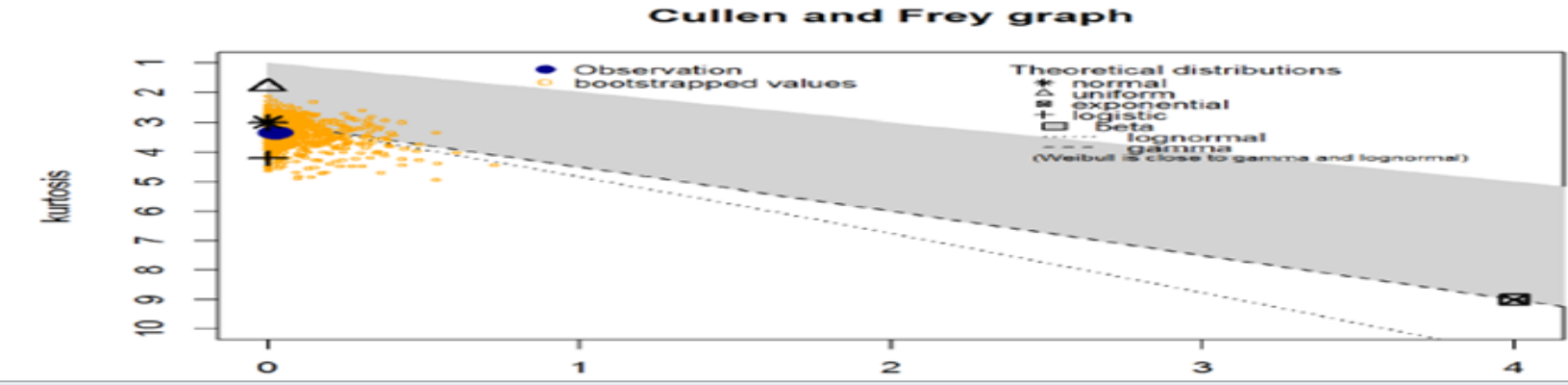


Empirical Distribution of Cluster 1

External Debt Ratio Distribution



Total Debt Ratio Distribution



Scatter Plot of Cluster 1

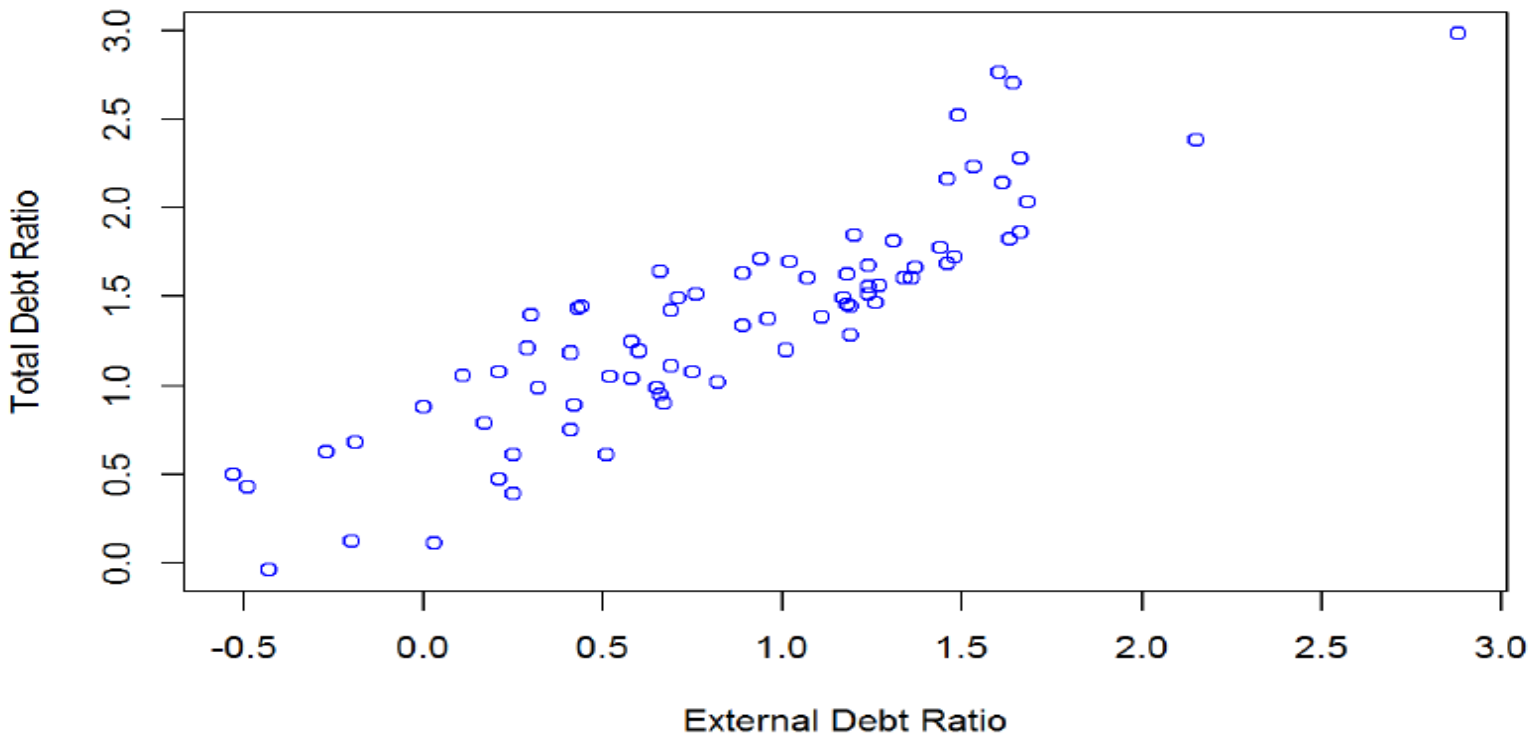
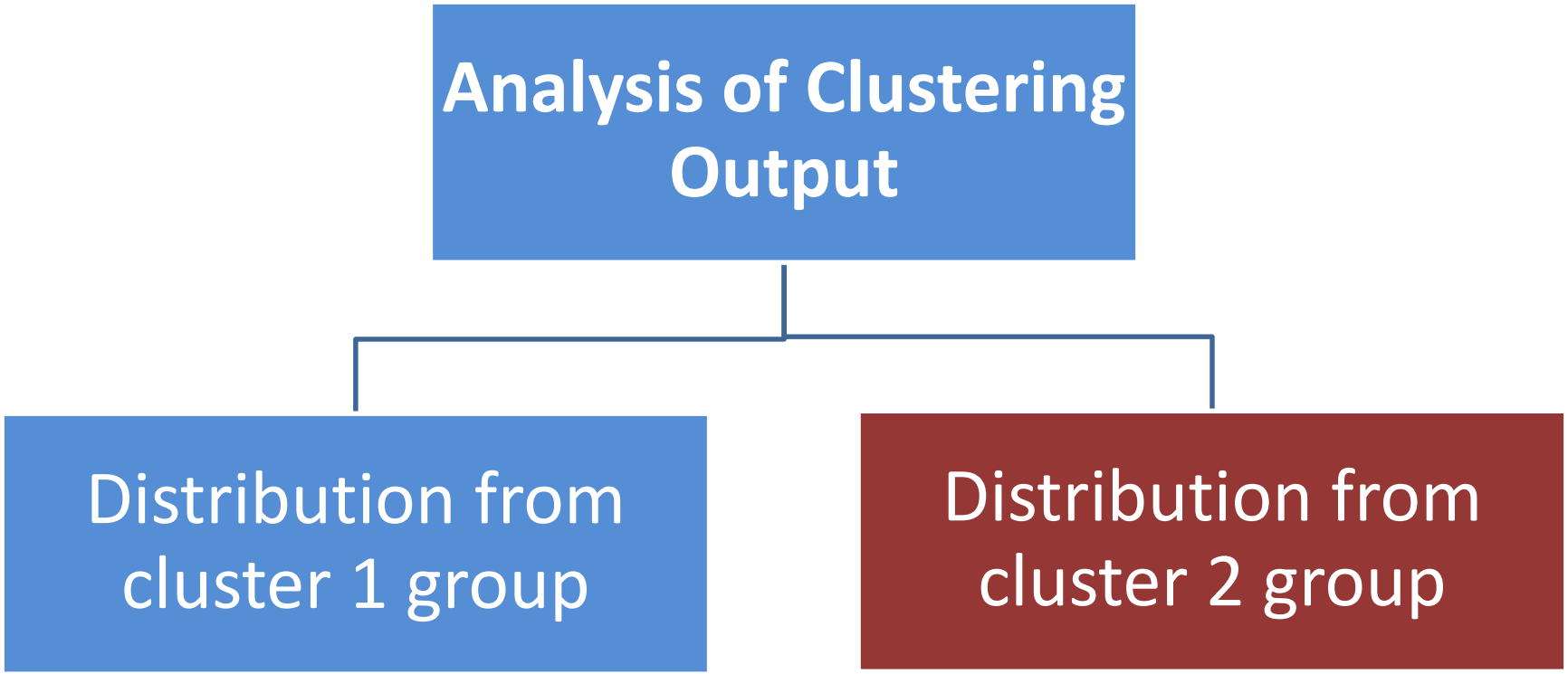


Figure 15. Scatter plot of external debt ratio and total debt ratio from cluster 1

Empirical Copula of Cluster 1

Figure 16. Plot of empirical copula from cluster 1


Analysis of Clustering Output
Distribution from cluster 1 group
Distribution from cluster 2 group

Density Plot of Cluster 2

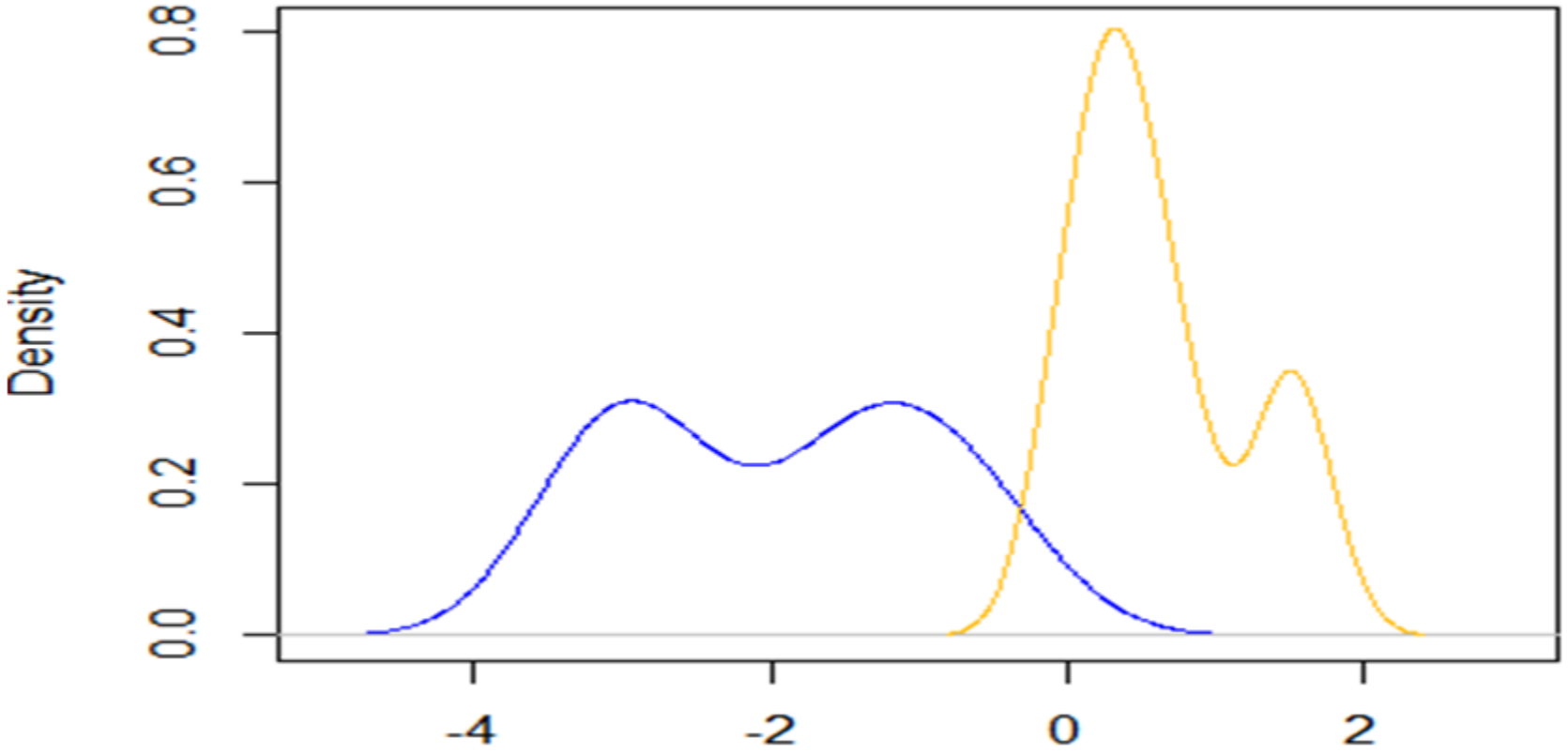


Figure 17. Density plot of external debt ratio and total debt ratio from cluster 2 group

Empirical Distribution of Cluster 2

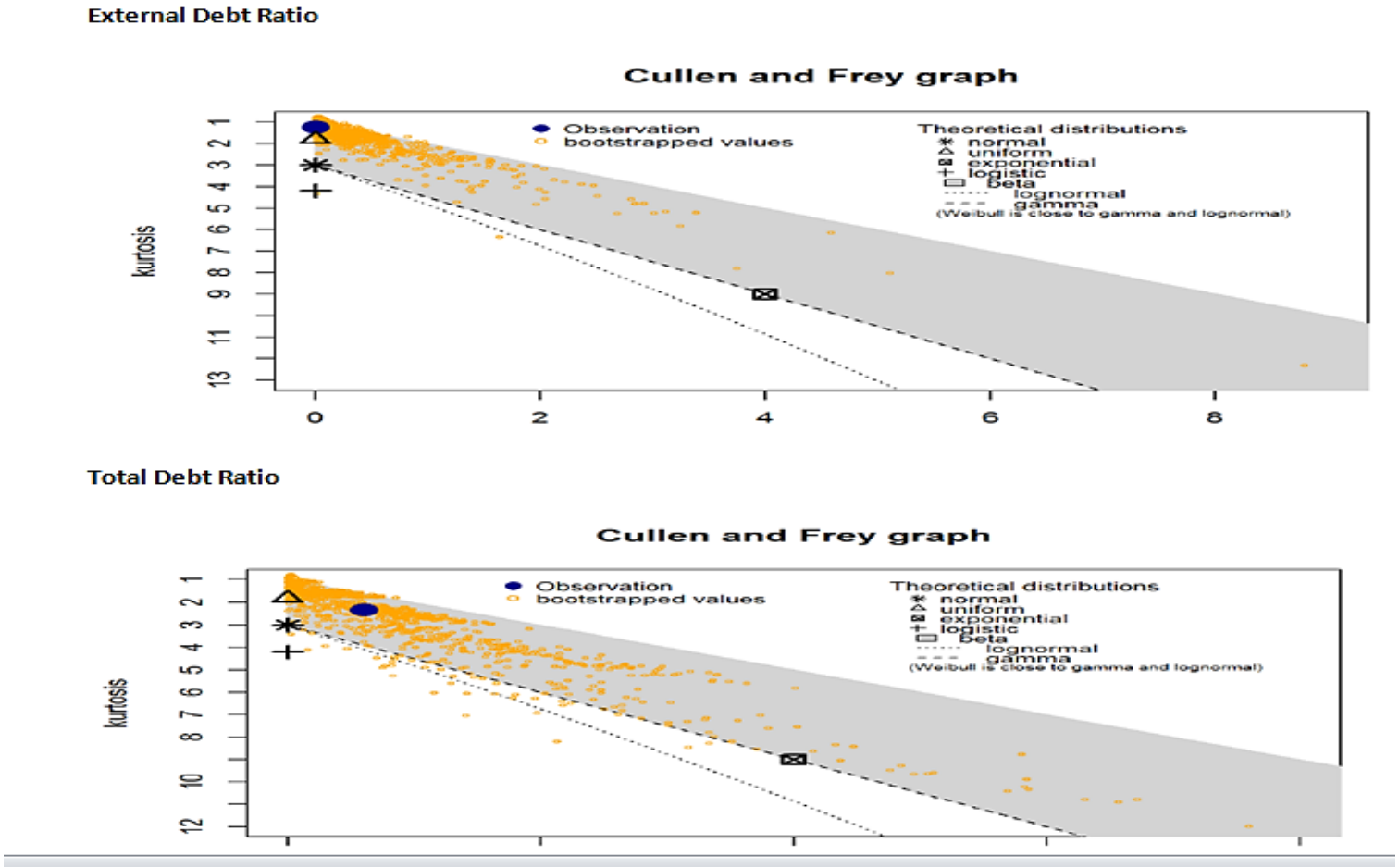
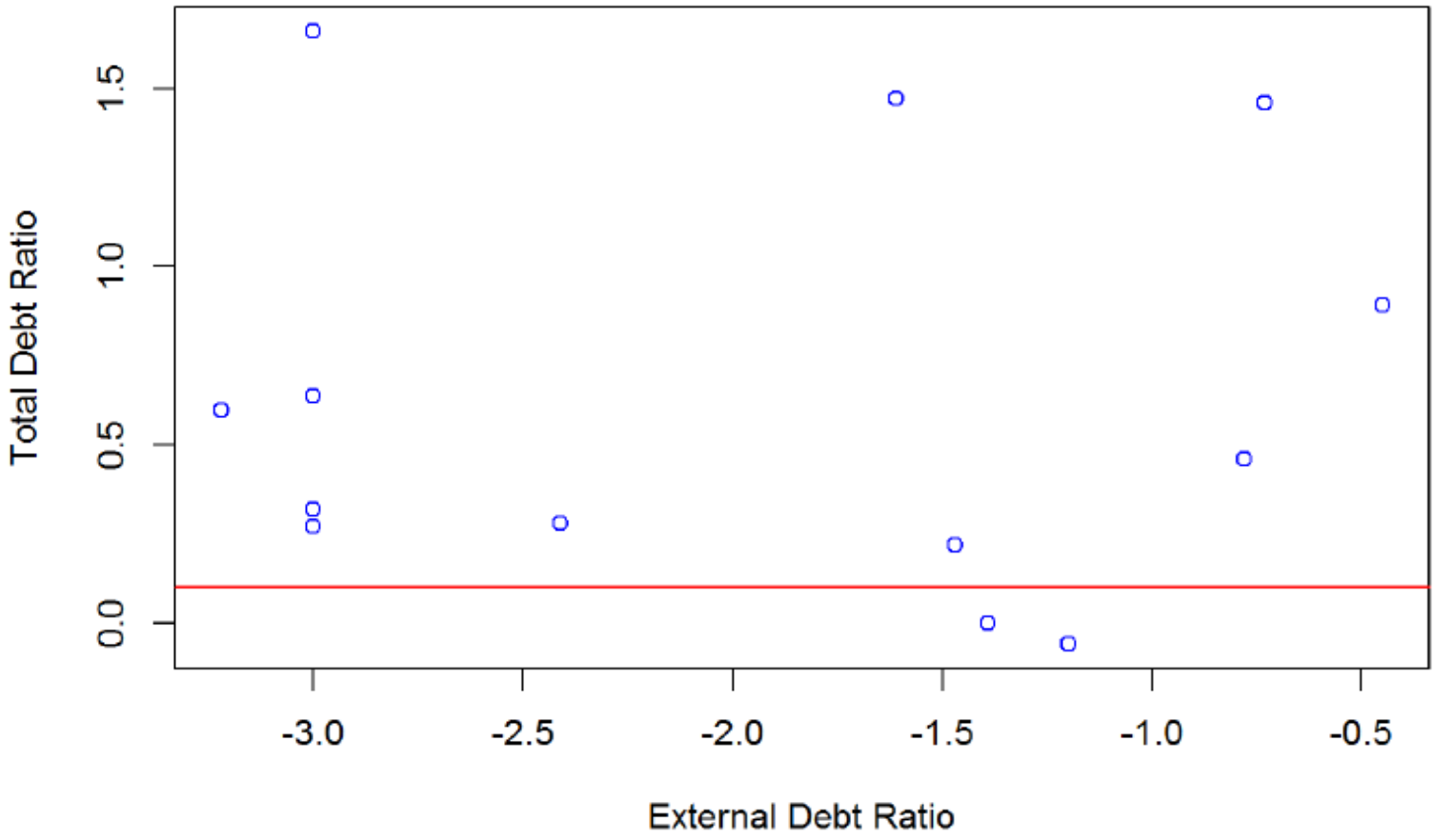


Figure 18. Empirical distribution check from cluster 2 group

Scatter Plot of Cluster 2



T. debt ratio is high E. debt ratio is small /Figure 4

Figure 19. Scatter plot of external debt ratio and total debt ratio from cluster 2

Empirical Copula of Cluster 2 (A)

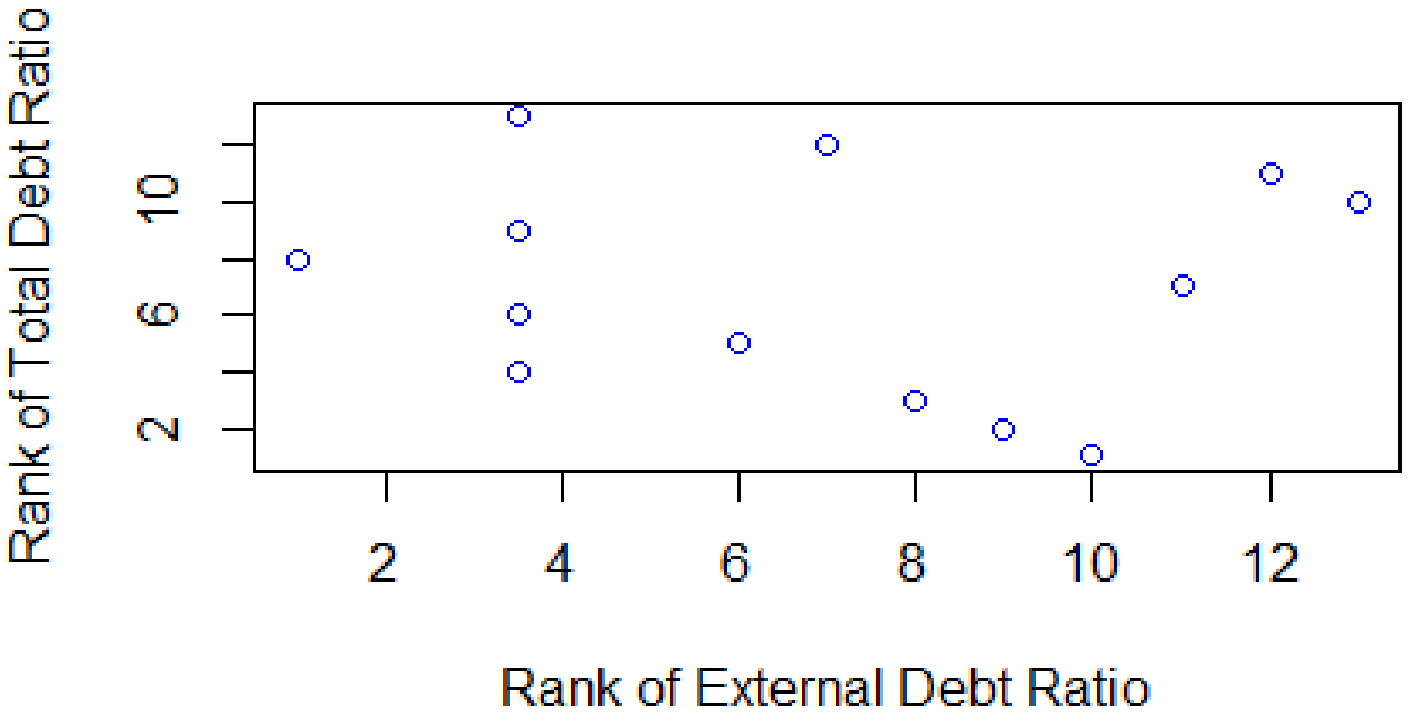


Figure 20. Plot of empirical copula from cluster 2

Empirical Copula of Cluster 2 (B)

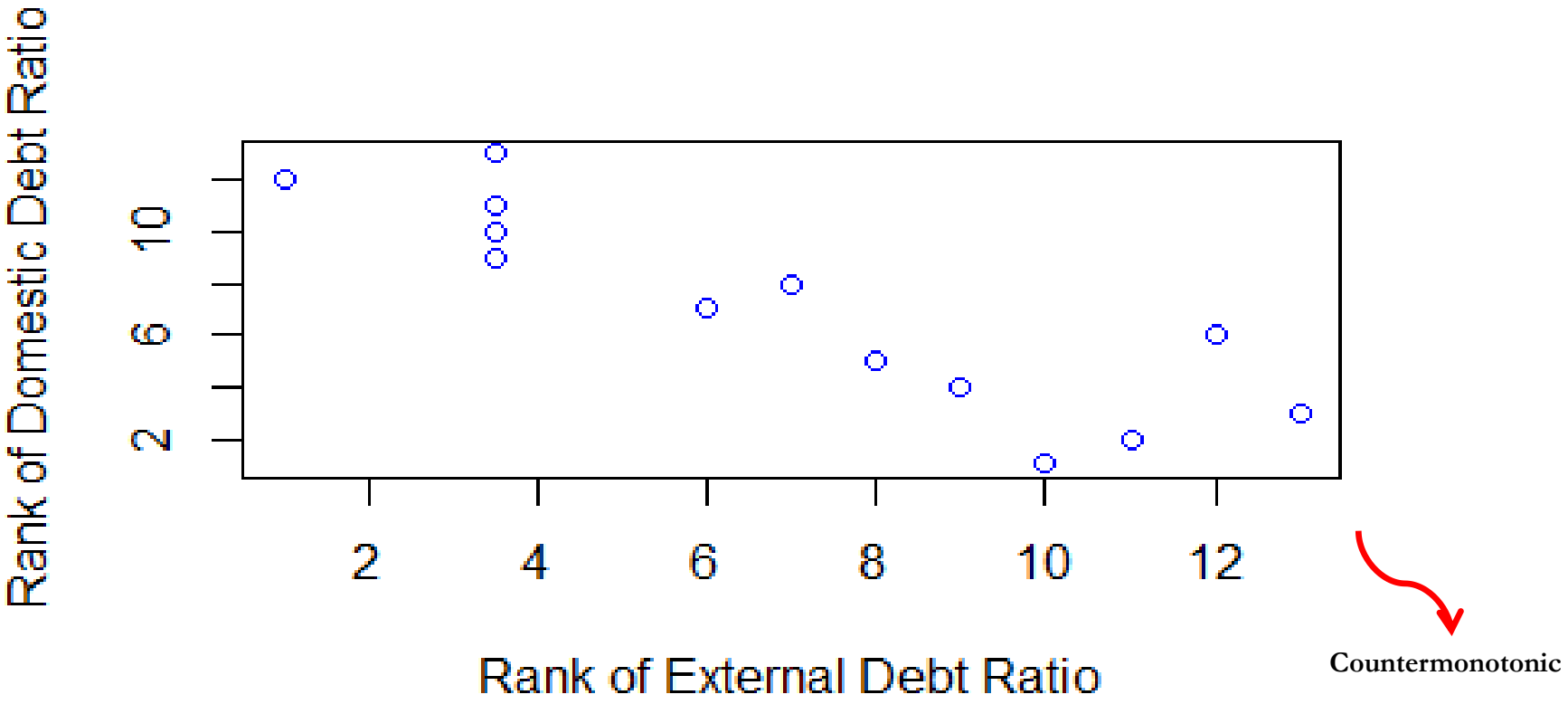


*Figure 21.*Plot of empirical copula between external debt ratio and domestic debt ratio from cluster 2

Clustering Output Based upon Timeline

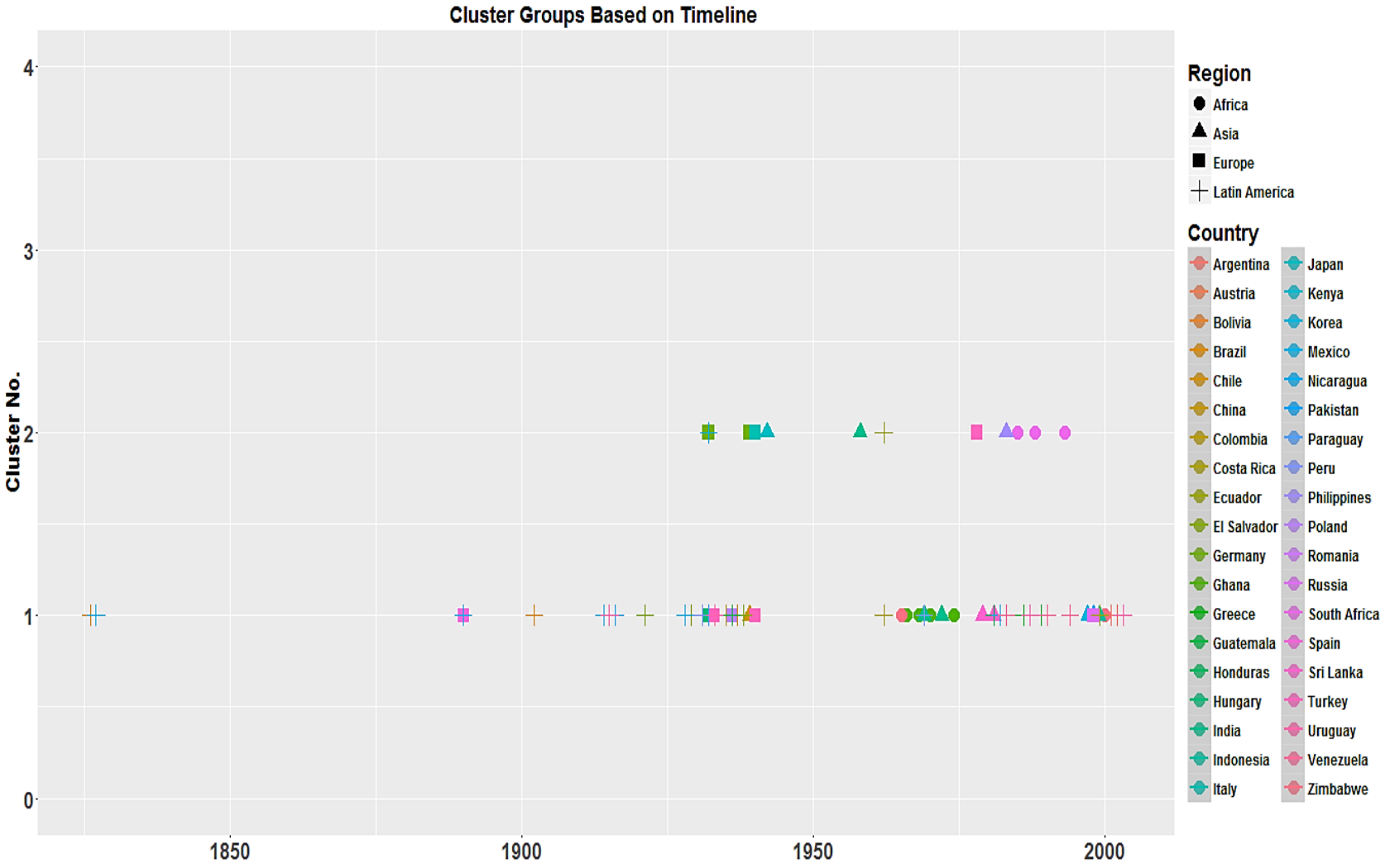


Figure 22. Plot of default countries based on timeline

Frequency of Default Countries

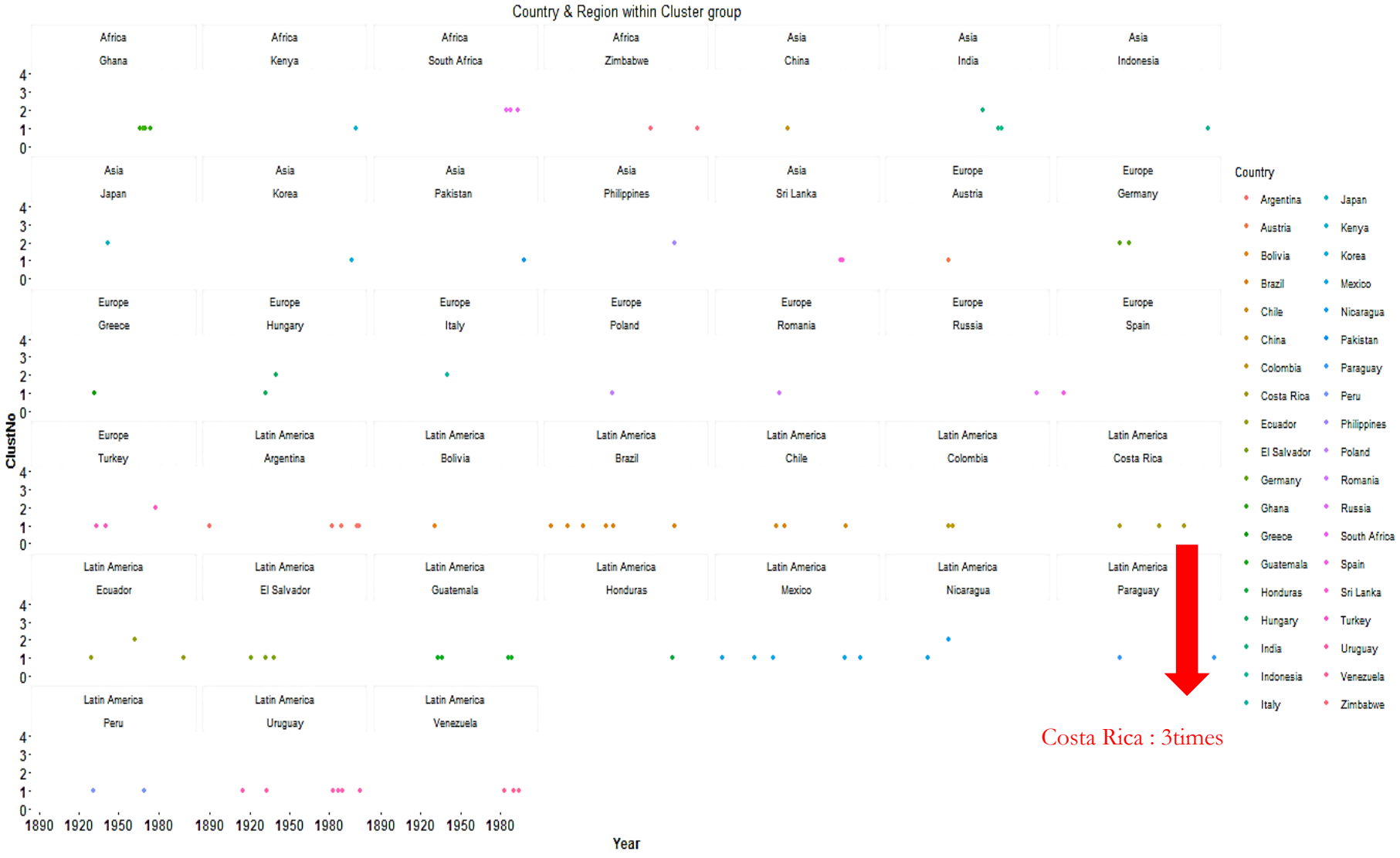


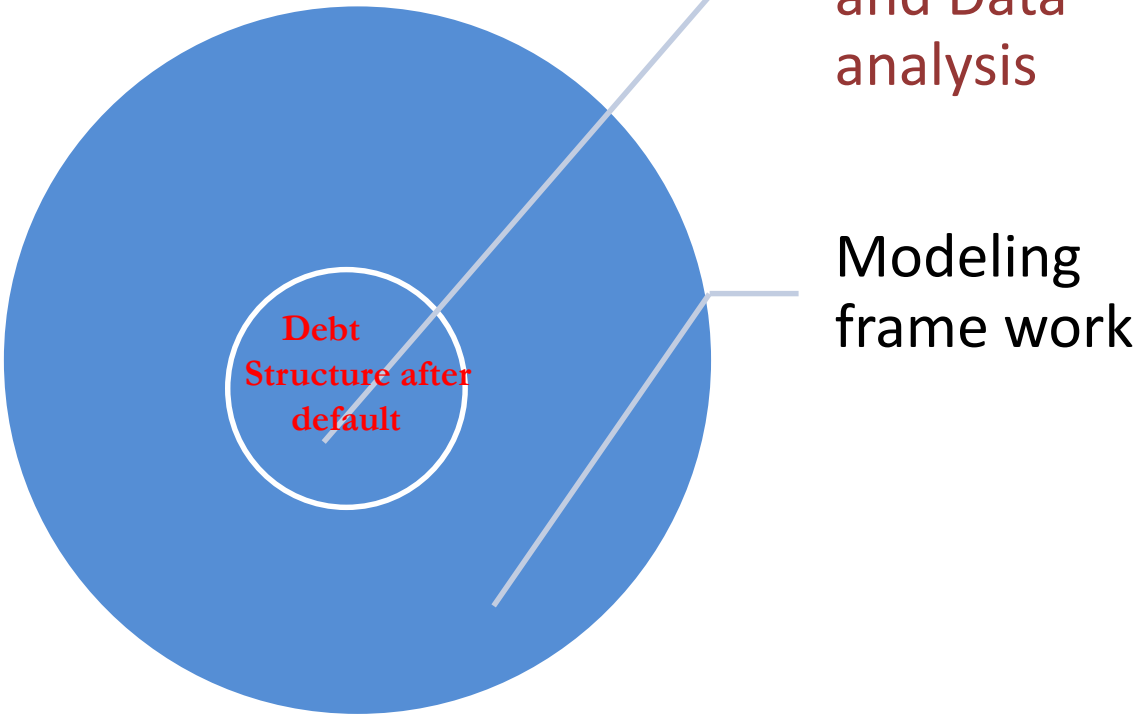
Figure 23. Plot of the number of default countries within each region

Methodology

Analysis of Debt Structure at Default

Analysis of government strategy after default

Analysis of Government Strategy After Default



Data Source and Data Analysis

Ten-year domestic debt data after government external default was selected

$$\text{Domestic debt adjusted for inflation} = \text{Debt at time } t+1/(1+f)$$

Default year is at time t. where f is the rate of inflation during year t

Modeling Framework

Function xyplot() of R software

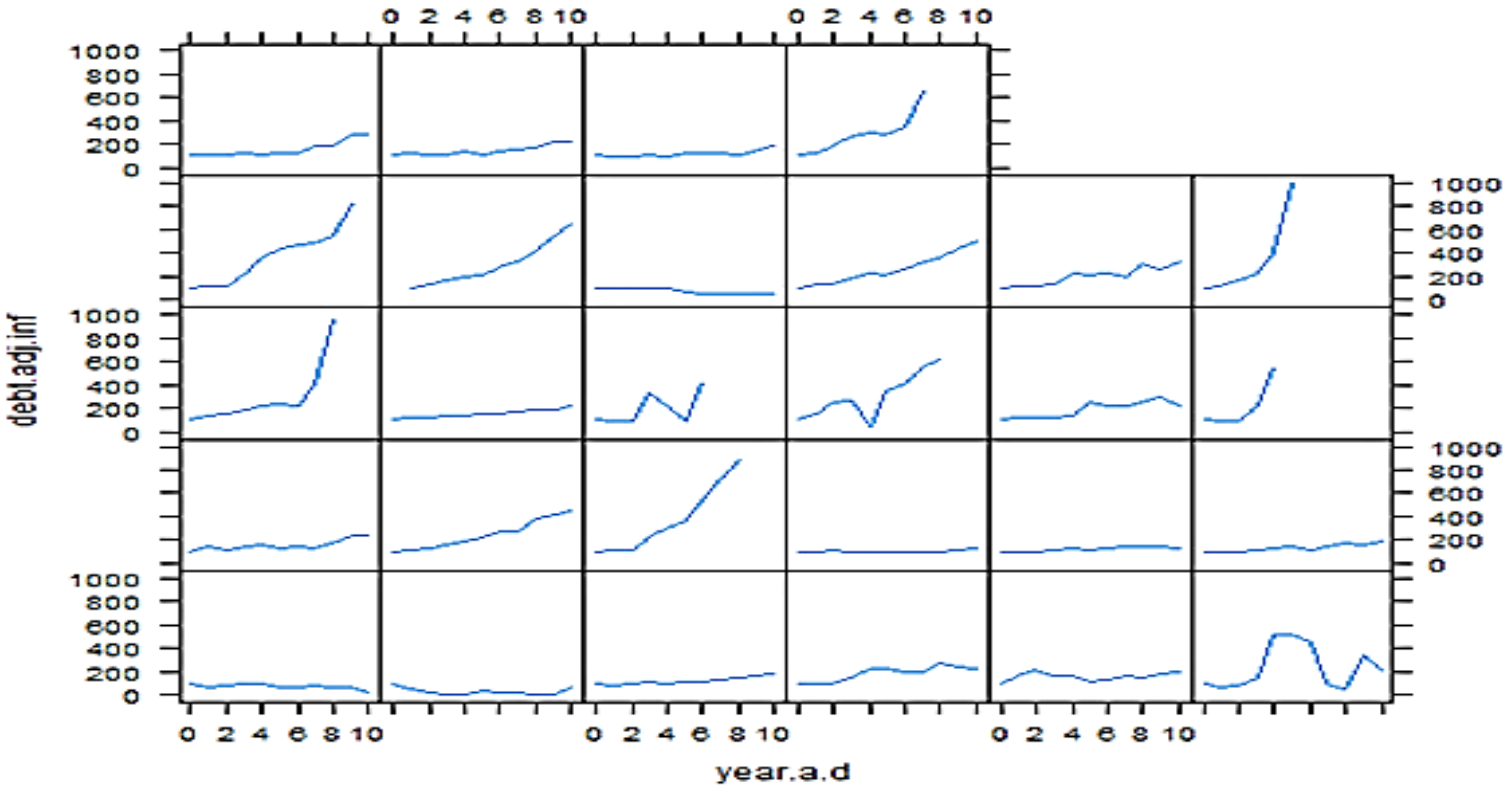
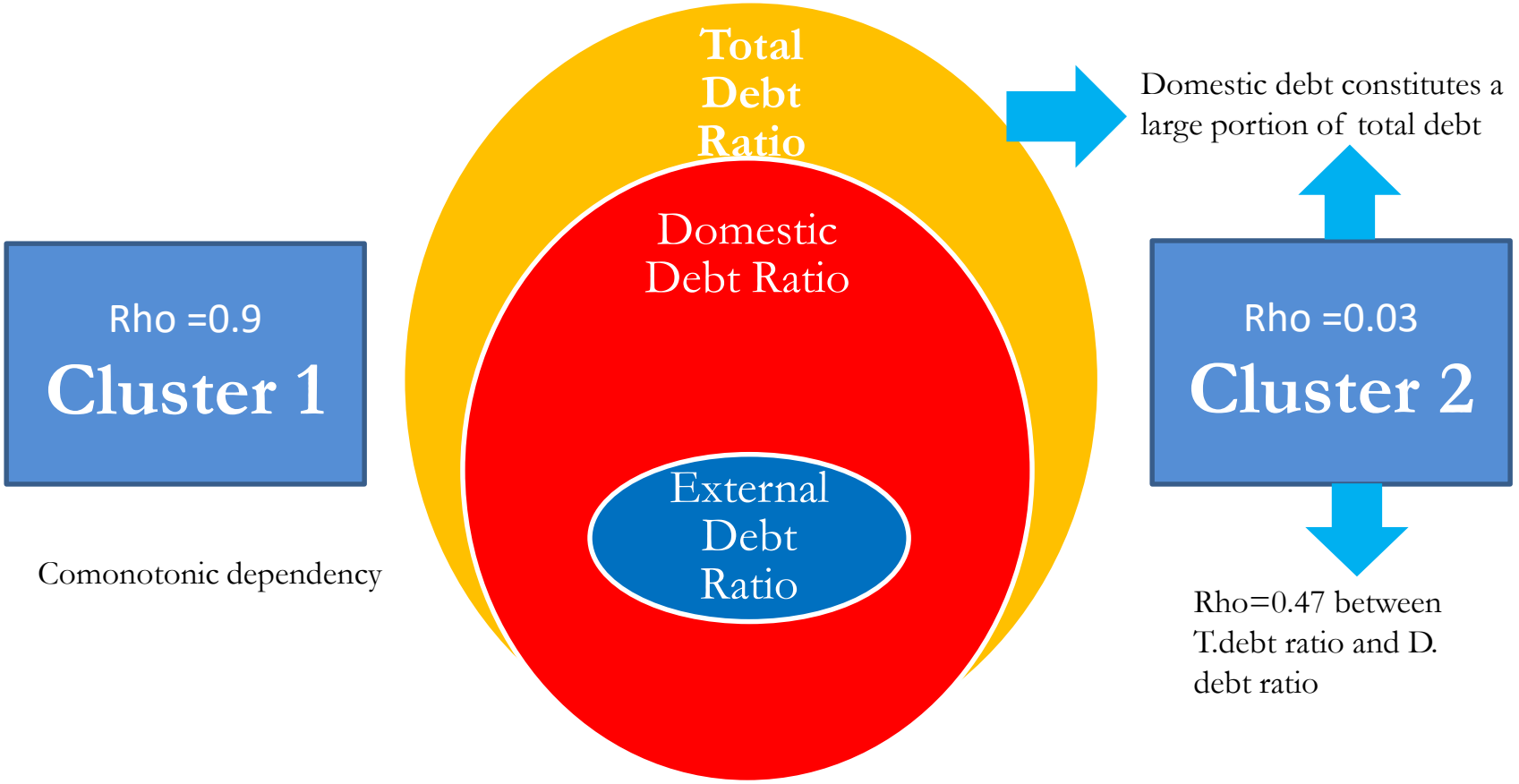


Figure 24. Adjustment of domestic debt over 10 year's period

Findings

Results of Descriptive Analyses + Modeling Results

Results of Descriptive Analyses



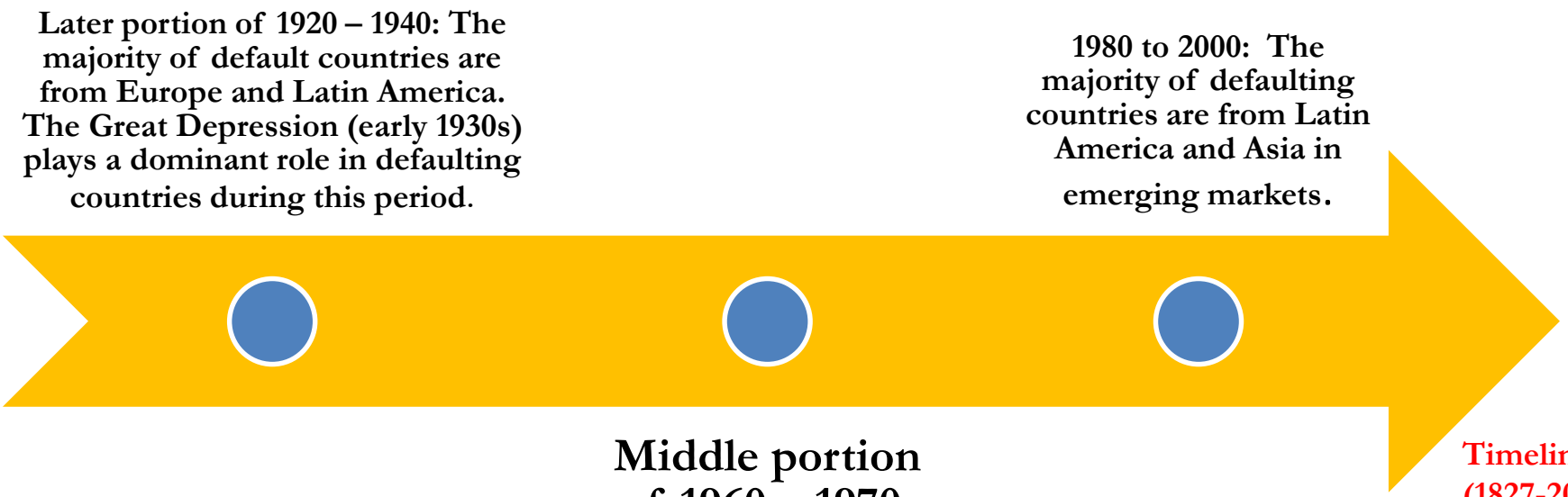
Modeling Results

- Examine the consistency of the serial defaults throughout history
- Examine authors' point of view that domestic debt plays a significant role
- Domestic debt is overlooked

Examine the Consistency of the Serial Defaults throughout History

Region	Country	Year	External.Debt.Ratio	Total.Debt.Ratio	Cluster.Num
Africa	Ghana	1966	2.13	4.51	1
	Ghana	1968	1.99	4.13	1
	Ghana	1970	1.5	3.25	1
	Ghana	1974	1.12	2.9	1
	Kenya	2000	1.99	3.03	1
	Zimbabwe	1965	1	2.4	1
	Zimbabwe	2000	1.35	4.03	1
Latin America	Argentina	1890	4.42	12.46	1
	Argentina	1982	1.79	3.44	1
	Argentina	1989	17.79	19.61	1
	Argentina	2001	1.68	2.86	1
	Argentina	2002	5.34	7.64	1
	Bolivia	1931	8.62	10.79	1
	Brazil	1826	4.4	5.56	1
	Brazil	1898	3.7	6.1	1
	Brazil	1902	3.45	5.3	1
	Brazil	1914	4.3	8.68	1
	Brazil	1931	4.99	8.51	1
	Brazil	1937	2.56	5.51	1
	Brazil	1983	0.83	1.98	1
	Chile	1931	3.51	4.29	1

Three Spike Episodes of Default in Countries Across Each Region

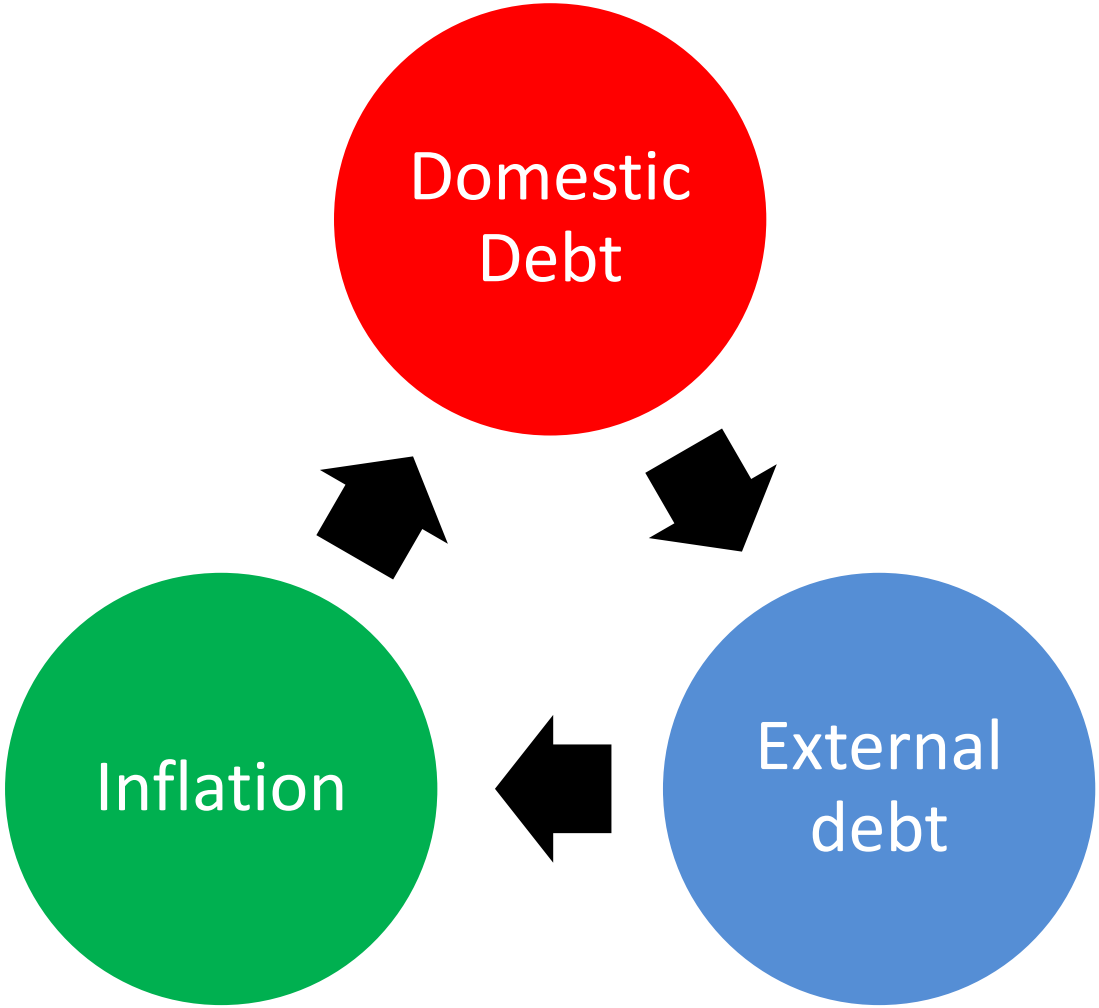


Cluster 1

Examine Authors' Point of View: Domestic Debt Plays a Significant Role

Region	Country	Default Year	External.Debt. Ratio	Total. Debt. Ratio	Cluster.Num
Africa	South Africa	1985	0.09	1.32	2
	South Africa	1988	0.05	1.38	2
	South Africa	1993	0.05	1.9	2
Asia	India	1958	0.2	4.35	2
	Japan	1942	0.04	1.83	2
	Philippines	1983	0.23	1.25	2
Europe	Germany	1939	0.05	1.31	2
	Hungary	1940	0.3	0.94	2
	Italy	1940	0.05	5.25	2
	Turkey	1978	0.25	1	2

Domestic Debt is Overlooked



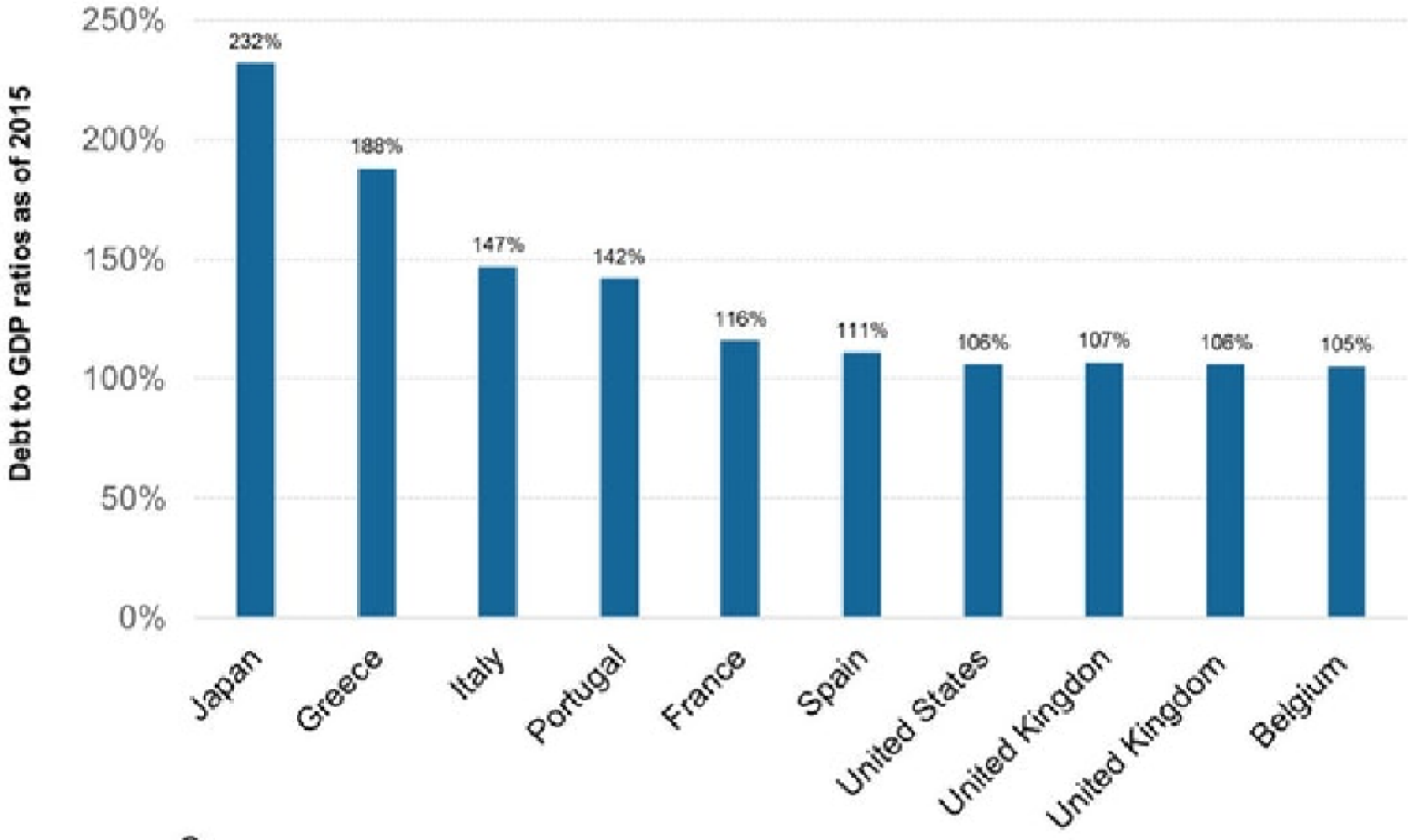
Conclusions

Is "this time really different"?

"The current boom, unlike the many booms that preceded catastrophic collapses in the past (even in our country) is built on sound fundamentals…."

Reinhart & Rogoff

Debt-To-GDP Ratio Of Various Economies



Market Realist

Source U.S.Department of the Treasury, Bureau of the Fiscal Service

Federal debt held by the public as a percentage of gross domestic product, from 1790 to 2013, projected to 2038