Does democracy favor a free market economy?

—— Principal component analysis and PAM algorithm in application

Zhiyi Jin

The Question

What are the economic foundations of prominent democracies? In particular, what kind of economic regulations or policies can promote democratic and open societies? One answer to this question may be the economic freedom brought by the free market. The free market, as previewed as the invisible hand by Adam Smith, is a type of economic system based on supply and demand without any intervention from the government. And it has been considered by some proponents of neoliberalism that the free market is a crucial factor for the quality of democracy and the failure of the free market could lead to the failure of a democratic society (Fracchiolla, 2018). As Fridman (1964) explained, in a free market system, citizens do not rely on the government for economic livelihood thus have the free will to oppose government policies. However, some people are skeptical about this view and claim that the free market could jeopardize both the economy and democracy due to unequal wealth distribution (Hahnel, 2009). This project will try to explore the question: "Does free market can promote political freedom or democracy?" or, in other words, "Does a democracy favor the free market system?"

Data Descriptions

The data set of *MeasuringSocialism* collected and complied by Joseph Nathan Cohen and Joseph van der Naald will be used in this project. Observations of the dataset cover 46 countries in the world from 2015 to 2017. And the data set measures various approaches by which governments intervene in the market including the government's expenditures, investment, tax revenue, and economic regulations. There is also one indicator `polity2` measuring the societal conformity to democratic ideals.

There are some redundant variables and missing values of the observations. Thus, before further analysis, the following steps are taken to clean the data set:
    (1) remove repeated indicators such as `ccode`, `scode`, and `exp.military.wdi`;
    (2) exclude indicators capturing minor types of tax sources and regulation terms with little variations in the observations;
    (3) only contain indicators with less than one-third missing values in the observations;
    (4) remove observations with missing values in `country`;
    (5) transform left missing values into 0.
Finally, the data contains 43 observations described by 85 key variables.

The main aim of this project is to identify the relationship between the free market and democracy, but I start by generating simple descriptive summaries for each of the variables measuring the free market (see Appendix Tab.1). These statistics capture the univariate distribution of each category. It can be seen that most variables range from 0 to 6, while some variables such as `gfinwort`, `soc.total` and `tax.all.tot` have large variances. This feature implies that it's sensible to conduct appropriates standardization of the data in further analysis, otherwise results would be dominated by few variables with large variance. With regard to the indicator of democracy, Tab.2 (see Appendix) shows that majority of countries are ranked high in the score with only a few exceptions such as Russia, Turkey, and China. A correlation matrix of the data also shows that there are significant correlations between some variables suggesting that some simplification of the free market index through dimension reduction should be possible.

Selection of Methods

Two methods will be applied in the project. The first one is principal component analysis, which is a technique from dimension reduction underlying the principles of eigenvalue to truncate a number of correlated variables into a smaller number of linearly combined variables called principal component (Mishra et al., 2017). This method is suitable for this project mainly because the original dataset consists of a large number of variables, most of which are interrelated. In order to identify patterns concerning the free market and to present the country with different levels of democracy in a way to highlight their difference and similarity in the market features, principal component analysis is an ideal choice. In addition, by reducing the dimensionality of the data, further analysis such as regression could be possible by incorporating those reduced sets of features as input.

The second method is PAM(partition around medoids) clustering. Clustering is a technique that can be used to divide data into different groups, where each group is similar to one another in certain ways (Bruce et al., 2020). The goal of clustering in this project is to identify meaningful categories of countries in the data. These groups could, on the one hand, validate the findings from the principal component analysis. On the other hand, it could be used as an outcome variable in a predictive regression to evaluate the influence of market features. PAM algorithm will be specifically applied mainly because it can minimize the dissimilarity of observations to their closest selected observations without the influence of outliers. The quality of clustering is maintained while less sensitive to extreme values.

Results

(1) Principal component analysis

Fig.1 (see Appendix) shows that the first component explains the largest part of the total variance with the amount of 39%. The second principle accounts for 9% of the variance. The first two principal components account for around 50% of the total variance. After that, there seems to be a decrease in the downward slope implying

that the variables of free-market features can be summarized by the first two components.

Then it can be seen from Fig.2 that the first component has a positive correlation with the majority of variables such as `reg.statecontrol`, `exp.consumption`, and `reg.prodmarket` but negatively correlated with few variables such as `reg.ret.pricecon`, `reg.ret.reglic` and `gfinworth`. And such it seems that it reflects a difference between a low level of market intervention and a high level of market intervention. However, it might be difficult to try to label this component, because as Fig.3 presented, variables are above the expected average contribution(red line) and the contribution of variables on the first component is somewhat uniform. It would be difficult to consider the most weighted variable in contributing to the first component.
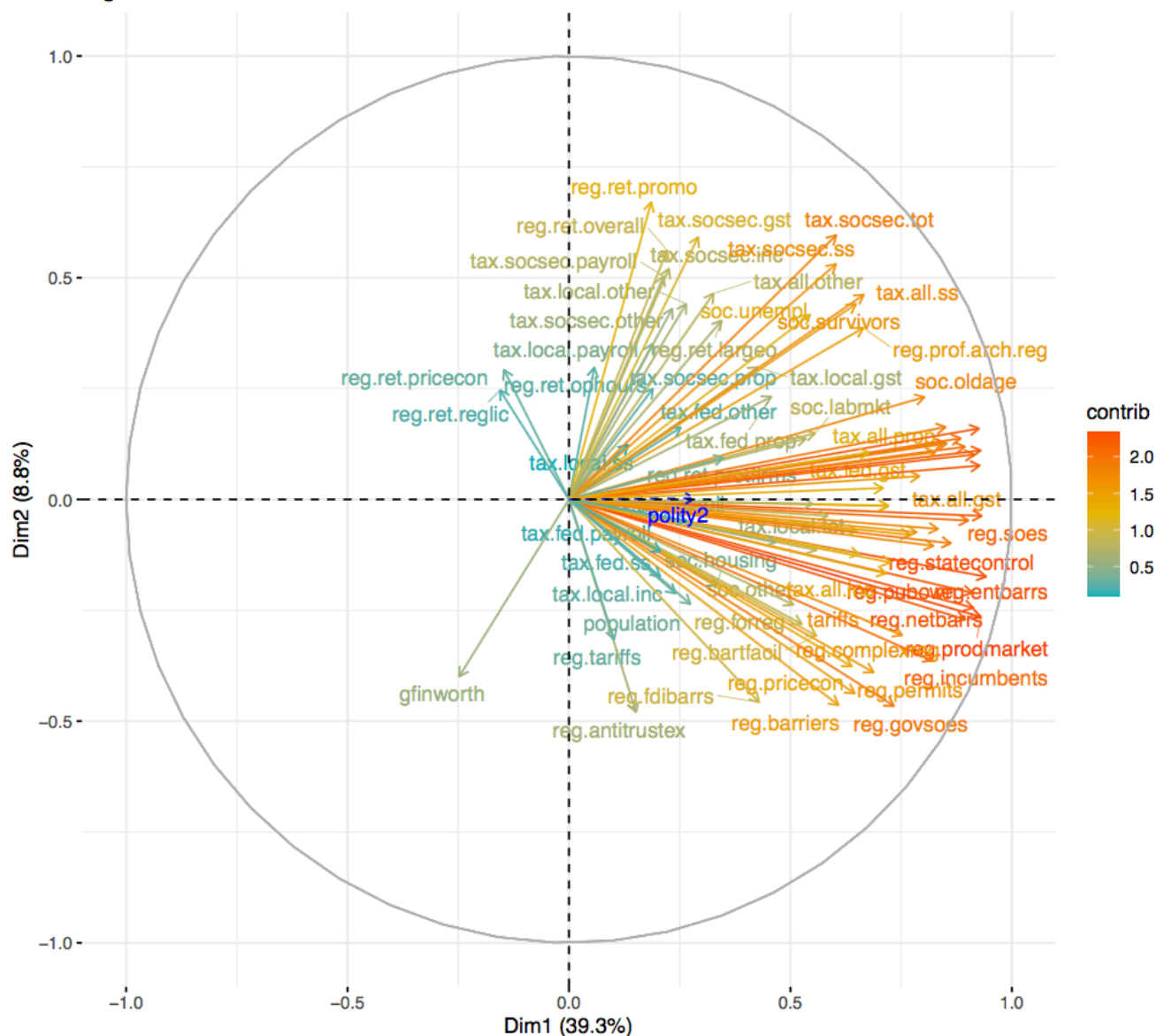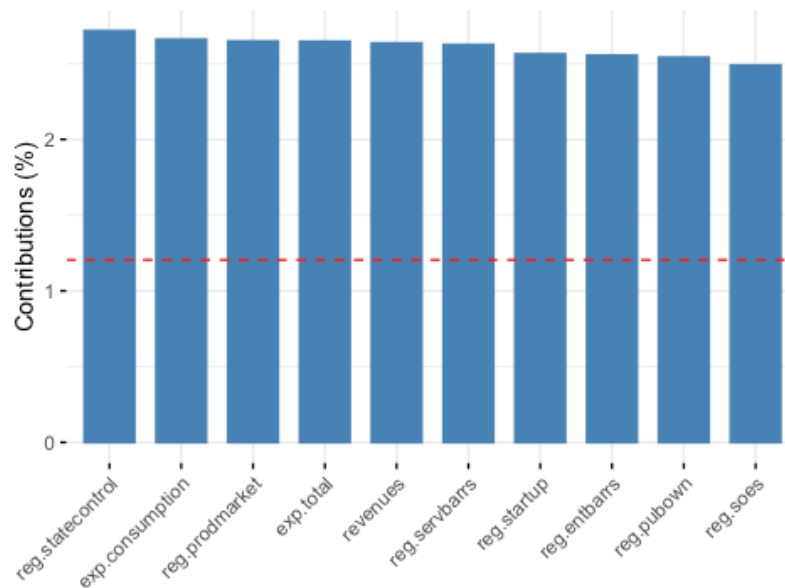

Fig.2 Variables – PCA

Fig.3 Contribution of variables to PC1

The second principal component is positively correlated to some tax-related variables such as `tax.socsec.gst`, `tax.socsec.tot` and `reg.ret.overall`, but negatively correlated with variables such as `reg.barriers` and `reg.tariffs`. It tempting to conclude that it tends to measure the difference in tax policies and regulation in the retail sector. Fig.4 further shows that `reg.ret.promo`, or the regulation on the retail sector could be the most contributed variables in the second component.

Fig.4 Contribution of variables to PC2
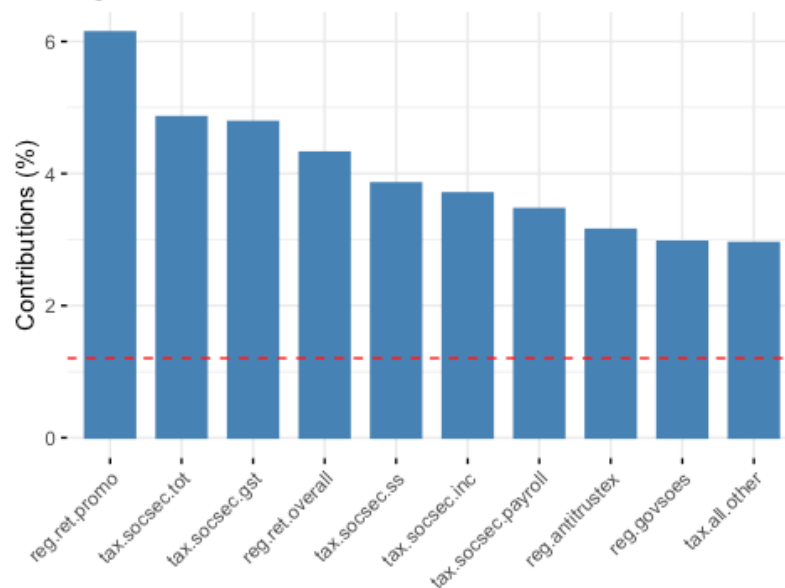
The results for each country are present in Fig.5 and colored by `polity2` as the supplementary variable. Countries seem to be divided into two groups in the plot. The large group lies on the right side around the center of coordinates. The small group lies on the far left side of the coordinates. Most countries with high scores on democracy concentrate on the right group, while countries with lower scores such as Russia, and China crowds in the left group though there are few exceptions such as Turkey. But in general, there seems to be a differential between these two groups of countries. And most importantly, the right group of countries is more likely to have a higher level of government intervention on market and more strict regulation on the product market, while the left group of countries seems to have fewer regulations despite more government financial assets and pricing regulation on the retail sector. This result seems to contradict the assumption that democracy favors the free market. And it turns that those countries with higher scores on democracy tend to have a higher level of market interventions.

Fig.5 Individuals - PCA
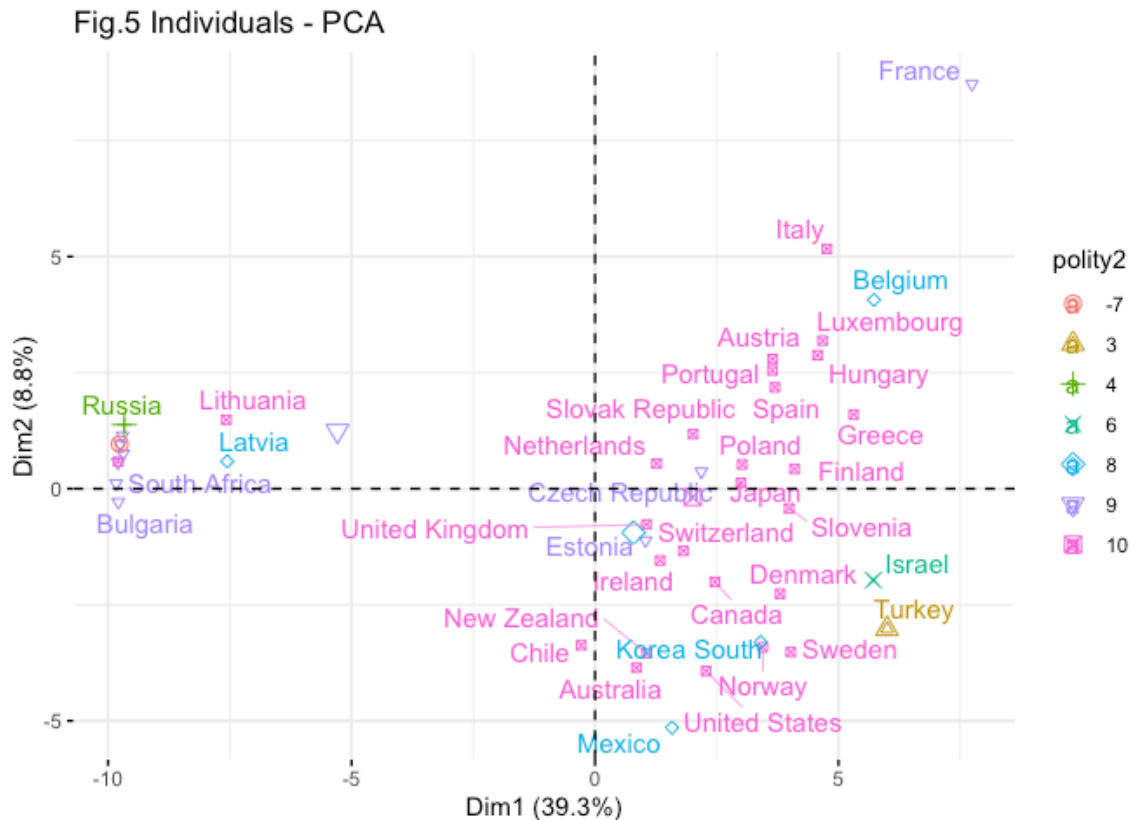
(Full display see Appendix Fig.5)

(2) PAM clustering

Clustering will be applied to testify the results from PCA and further explore the dataset. PAM algorithm needs to divide the dataset into clusters for a chosen k number in advance. Both silhouette and gap statistics indicate that the optimal number of clusters is two (see Appendix Fig.6). Fig.7 shows the result of two clusters, which are significantly similar to the result from principal component

analysis. Thus, two groups can be differentiated by labeling the larger one as "democratic group" and the smaller one as "autocratic group". The validity of the clusters is evaluated in Fig.8 (see Appendix), which shows that clusters are well-located without any values below 0.

Fig.7 Cluster plot
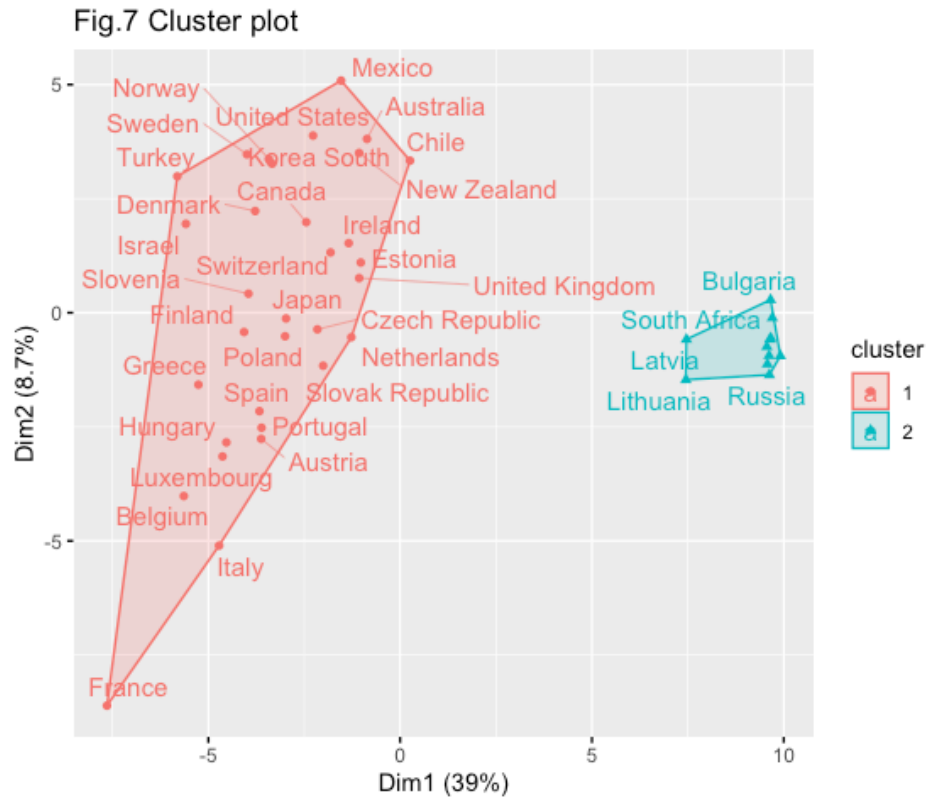
(Full display see Appendix Fig.7)

Fig.9 shows the analysis of box plots for each group on several key variables of the free-market features that are drawn from the principal component analysis. Interestingly, there is actually no significant difference in their regulation of the retail sector. And because of the small group size of the autocratic group and a large amount of missing data, there is no way to tell the difference on key variables such as the government's expenditures and tax barriers. This result proves that the conclusion drawn in the PCA analysis might have been misled by corrupted data with a strong disturbance of missing values. Therefore, the previous conclusion that democratic countries do not favor the free market can't be supported, nor does the opposite argument that democracy favors the free market. However, it would be still interesting to think about the division of countries and the reasons behind those missing values. Missing records of those indexes could be caused by the lack of transparency of the government or some other bureaucratic issues, which are usually more common in countries with a lower level of political freedom. It might not be a coincidence that democratic countries have more records on their intervention of the market and lead to a falsification of PCA's result.

Fig.9 Result Analysis - PAM
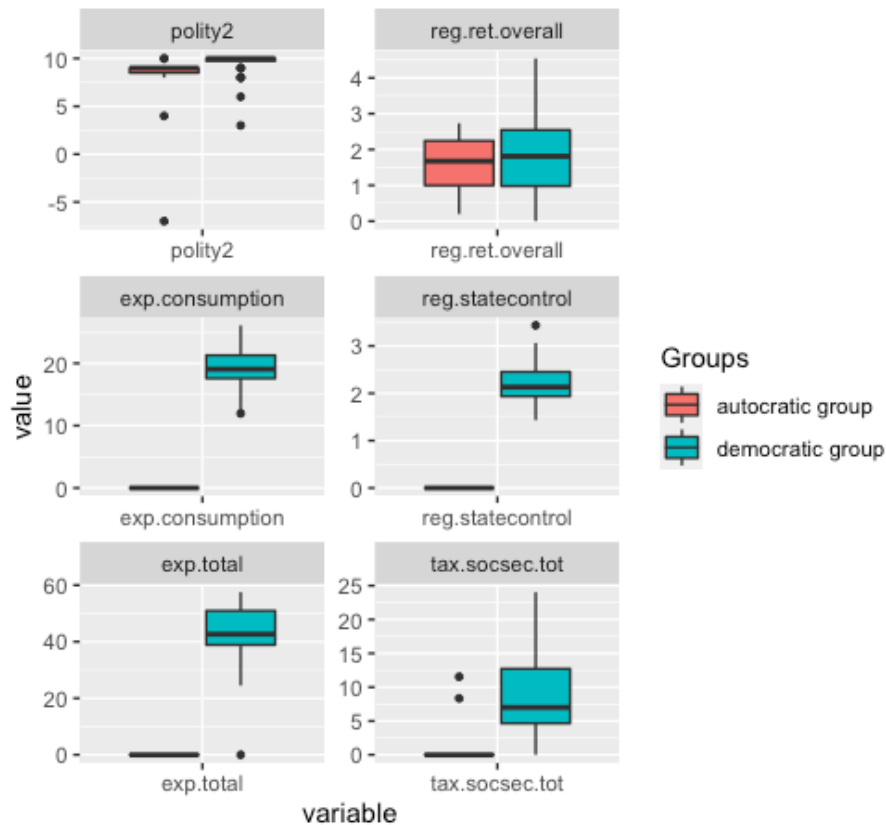
Limitations

This project tends to explore the relationship between the free market and democracy by the application of principal component analysis and PAM algorithm to the $MeasuringSocialism$ dataset. Considering the complexity of this question and the simplicity of this project, there are certainly many challenges. Here some of them are listed. First of all, the result of PCA can not be validated by the analysis of PAM clustering. PCA does not seem to provide meaningful results in this project. The distinctiveness of the eigenvalues is a fundamental assumption of PCA, yet the principal components in the project are difficult to interpret or seem to lack meanings due to the high dimensions of the original dataset. Secondly, it was a hard time handling missing data and outliers in this dataset. Small subpopulations or those countries in the autocratic group seem not well represented in the dataset with too many unknown values on their market features. This definitely affects the robustness of PCA as well as the clustering algorithm. Thirdly, the method of PCA relies on the linear correlation between variables in the dataset. Nevertheless, some indexes in the $MeasuringSocliasm$ dataset are actually under the same categories but at different levels such as tax at local and state levels. If these variables are not linearly correlated, PCA might not be an appropriate approach. Overall, while these limitations do not invalidate the exploratory purpose of this simple project, further

improvements such as the imputation of missing data and the examination of assumptions in the context of PCA should be taken into consideration.

References

[1] Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python*. O'Reilly Media.

[2] Fracchiolla, D. (2018). Free Market and Democracy. *Sociology and Anthropology*, 6(2), 259–273. https://doi.org/10.13189/sa.2018.060207

[3] Friedman, M. (1964). *Capitalism and freedom*. Chicago: University of Chicago Press.

[4] Hahnel, R. (2009). Why the market subverts democracy. *American behavioral scientist*, 52(7), 1006-1022.

[5] Mishra, S. P., Sarkar, U., Taraphder, S., Datta, S., Swain, D. P., Saikhom, R., & Laishram, M. (2017). Multivariate statistical data analysis-principal component analysis (PCA). *International Journal of Livestock Research*, 7(5), 60-78.

Appendix

Tab.1 Descriptive Statistics

	Min	Q1	Med	Mean	Q3	Max
exp.total	0.00	0.00	40.12	31.91	47.85	57.51
exp.consumption	0.00	5.98	18.17	14.18	20.05	26.10
ginv	0.00	0.00	3.00	2.55	4.07	5.29
revenues	0.00	0.00	38.56	30.31	44.80	55.02
gfinworth	-136.11	-63.45	-5.15	-21.90	0.00	247.09
tariffs	0.00	0.66	1.88	1.59	1.88	6.48
soc.oldage	0.00	0.00	6.42	5.91	9.52	14.45
soc.survivors	0.00	0.00	0.32	0.73	1.32	2.67
soc.incap	0.00	0.00	1.81	1.64	2.52	5.58
soc.health	0.00	0.00	5.30	4.56	6.59	14.05
soc.family	0.00	0.00	1.55	1.54	2.65	3.54
soc.labmkt	0.00	0.00	0.23	0.36	0.57	2.05
soc.unempl	0.00	0.00	0.33	0.57	0.89	2.95
soc.housing	0.00	0.00	0.10	0.24	0.35	1.53
soc.other	0.00	0.00	0.25	0.41	0.68	2.19
soc.total	0.00	3.84	19.81	16.44	24.64	32.22
reg.prodmarket	0.00	0.46	1.29	1.10	1.48	2.46
reg.statecontrol	0.00	0.72	2.02	1.65	2.22	3.44
reg.pubown	0.00	0.76	2.43	2.04	2.94	4.36
reg.bizinvolve	0.00	0.48	1.35	1.26	1.83	3.32
reg.entbarrs	0.00	0.57	1.55	1.25	1.75	2.78
reg.startup	0.00	0.46	1.45	1.39	2.15	3.08
reg.barriers	0.00	0.06	0.35	0.40	0.56	1.52
reg.soes	0.00	0.70	2.60	2.27	3.26	5.65
reg.commandreg	0.00	0.27	1.47	1.45	2.23	3.49
reg.pricecon	0.00	0.19	0.94	1.07	1.48	3.47
reg.permits	0.00	0.00	2.00	2.29	4.00	6.00
reg.complexreg	0.00	0.20	1.36	1.38	2.07	3.75
reg.commrules	0.00	0.00	0.26	0.46	0.89	1.50
reg.aburd.major	0.00	0.00	1.40	1.12	1.70	2.60
reg.aburd.sole	0.00	0.00	0.40	0.60	1.00	2.60
reg.legbarrs	0.00	0.10	0.67	0.70	1.18	1.97
reg.incumbents	0.00	0.32	1.15	1.00	1.38	2.72

reg.antitrustex	0.00	0.00	0.00	0.27	0.00	2.95
reg.tariffs	0.00	0.00	0.00	0.09	0.00	2.50
reg.directcontrols	0.00	0.21	0.96	1.07	1.67	3.30
reg.fdibarrs	0.00	0.01	0.20	0.32	0.46	1.74
reg.govnet	0.00	0.29	2.37	2.14	3.29	5.00
reg.netbarrs	0.00	0.48	2.47	2.04	2.97	3.95
reg.govsoes	0.00	0.38	3.00	2.66	4.12	5.62
reg.servbarrs	0.00	0.57	3.01	2.45	3.88	4.62
reg.forreg	0.00	0.00	0.34	0.44	0.57	1.84
reg.bartfacil	0.00	0.00	0.32	0.73	1.34	3.16
reg.prof.legal.reg	0.00	0.28	2.40	2.21	3.47	5.46
reg.prof.arch.reg	0.00	0.00	0.75	1.14	2.22	3.35
reg.prof.eng.reg	0.00	0.00	0.19	0.91	1.94	3.35
reg.prof.acc.reg	0.00	0.48	1.71	1.68	2.67	4.85
reg.ret.overall	0.00	0.96	1.79	1.79	2.51	4.54
reg.ret.reglic	0.00	1.84	3.60	3.30	4.80	6.00
reg.ret.largeo	0.00	0.00	0.00	1.91	4.00	6.00
reg.ret.protfirms	0.00	0.00	0.00	1.47	3.00	6.00
reg.ret.ophours	0.00	0.00	0.00	1.19	2.64	5.00
reg.ret.pricecon	0.00	0.86	1.71	1.65	2.49	5.14
reg.ret.promo	0.00	0.00	0.00	1.21	2.00	6.00
tax.fed.tot	0.00	11.24	17.70	16.02	23.60	33.67
tax.fed.inc	0.00	3.36	6.74	6.73	10.01	18.40
tax.fed.ss	0.00	0.00	0.00	0.71	0.35	10.44
tax.fed.payroll	0.00	0.00	0.00	0.22	0.01	4.62
tax.fed.prop	0.00	0.00	0.18	0.48	0.68	3.19
tax.fed.gst	0.00	3.70	9.68	7.86	11.65	14.86
tax.fed.other	-0.36	0.00	0.00	0.03	0.02	0.37
tax.local.tot	0.00	0.36	1.58	2.85	3.93	15.52
tax.local.inc	0.00	0.00	0.00	1.50	1.00	15.14
tax.local.ss	0.00	0.00	0.00	0.00	0.00	0.08
tax.local.payroll	0.00	0.00	0.00	0.03	0.00	0.86
tax.local.prop	0.00	0.19	0.64	0.87	1.38	3.26
tax.local.gst	0.00	0.00	0.06	0.36	0.44	2.09
tax.local.other	0.00	0.00	0.00	0.09	0.01	1.70
tax.socsec.tot	0.00	0.00	6.23	6.66	12.07	24.04

tax.socsec.inc	0.00	0.00	0.00	0.13	0.00	5.10
tax.socsec.ss	0.00	0.00	6.04	6.43	12.07	16.36
tax.socsec.payroll	0.00	0.00	0.00	0.02	0.00	0.92
tax.socsec.prop	0.00	0.00	0.00	0.00	0.00	0.08
tax.socsec.gst	0.00	0.00	0.00	0.07	0.00	1.63
tax.socsec.other	0.00	0.00	0.00	0.00	0.00	0.04
tax.all.tot	0.00	21.74	31.57	26.53	36.80	46.13
tax.all.inc	0.00	5.94	8.45	8.78	12.91	29.29
tax.all.ss	0.00	0.03	7.28	7.16	12.25	16.72
tax.all.payroll	0.00	0.00	0.00	0.33	0.10	4.62
tax.all.prop	0.00	0.32	1.28	1.50	2.70	4.19
tax.all.gst	0.00	6.06	10.81	8.56	11.85	16.99
tax.all.other	-0.36	0.00	0.00	0.12	0.06	1.70

Tab.2 Country Ranks of Democracy

country	polity2
United States	10
Canada	10
Chile	10
United Kingdom	10
Ireland	10
Netherlands	10
Luxembourg	10
Switzerland	10
Spain	10
Portugal	10
Poland	10
Austria	10
Hungary	10
Slovak Republic	10
Italy	10
Slovenia	10
Greece	10
Cyprus	10
Lithuania	10
Finland	10
Sweden	10
Norway	10
Denmark	10
Japan	10
Australia	10
New Zealand	10
France	9
Czech Republic	9
Croatia	9
Bulgaria	9
Romania	9
Estonia	9
South Africa	9
India	9

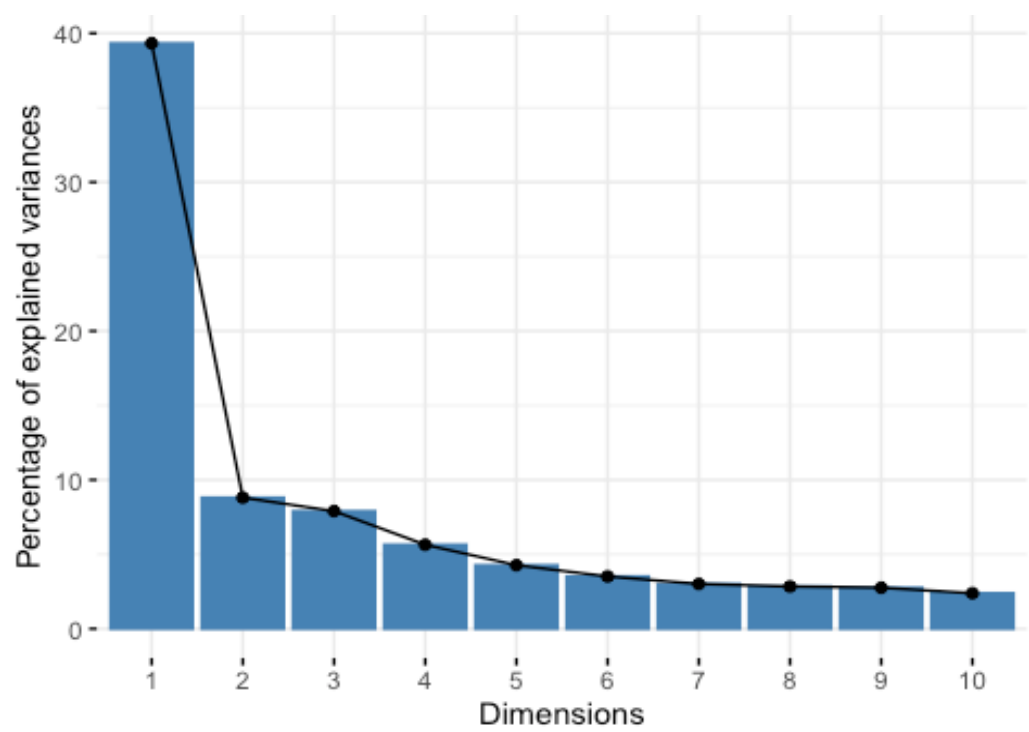|	
Indonesia	9
Mexico	8
Belgium	8
Latvia	8
Korea South	8
Israel	6
Russia	4
Turkey	3
China	-7

Fig.1 Scree Plot
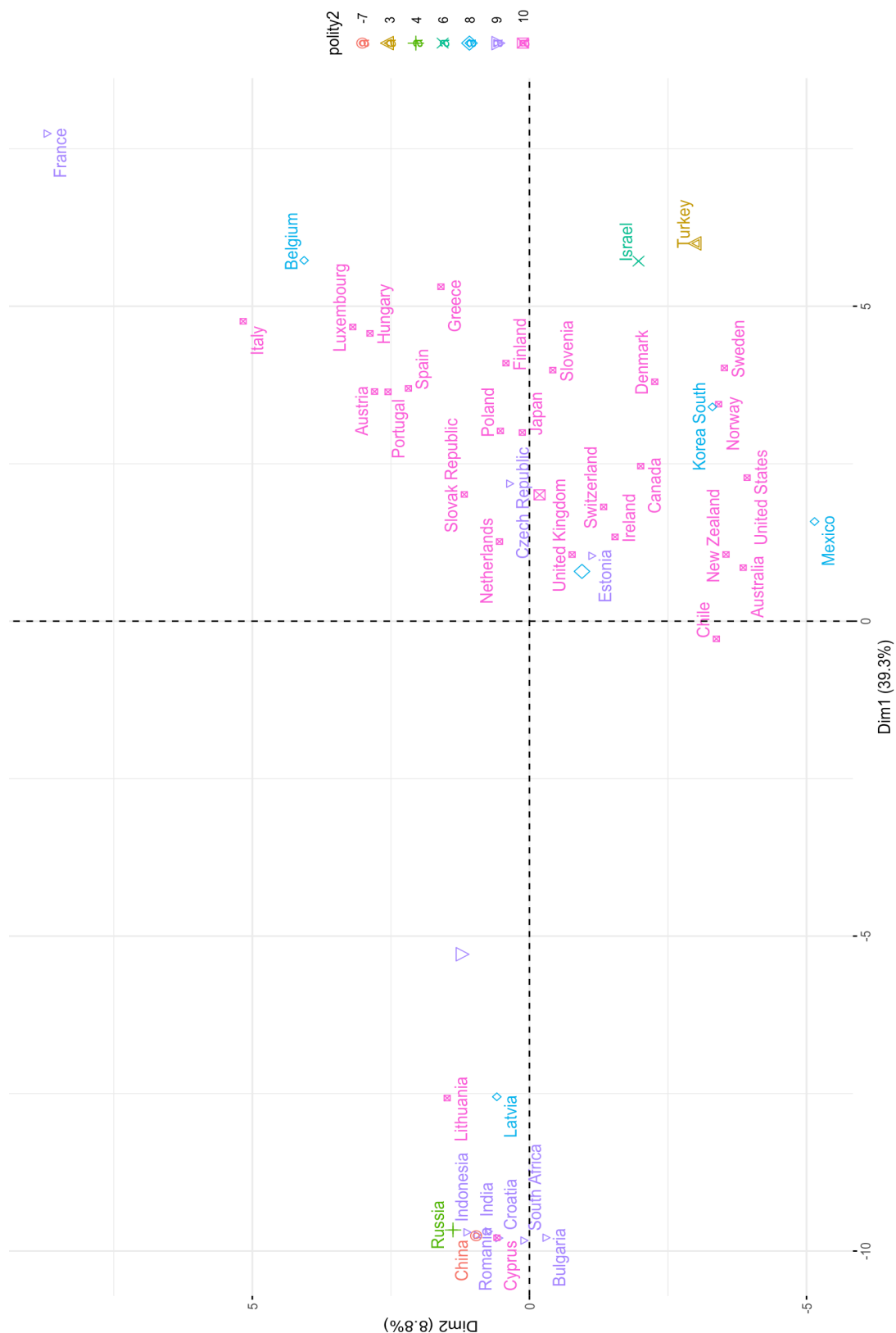
Fig. 5 Full display of Individuals – PCA

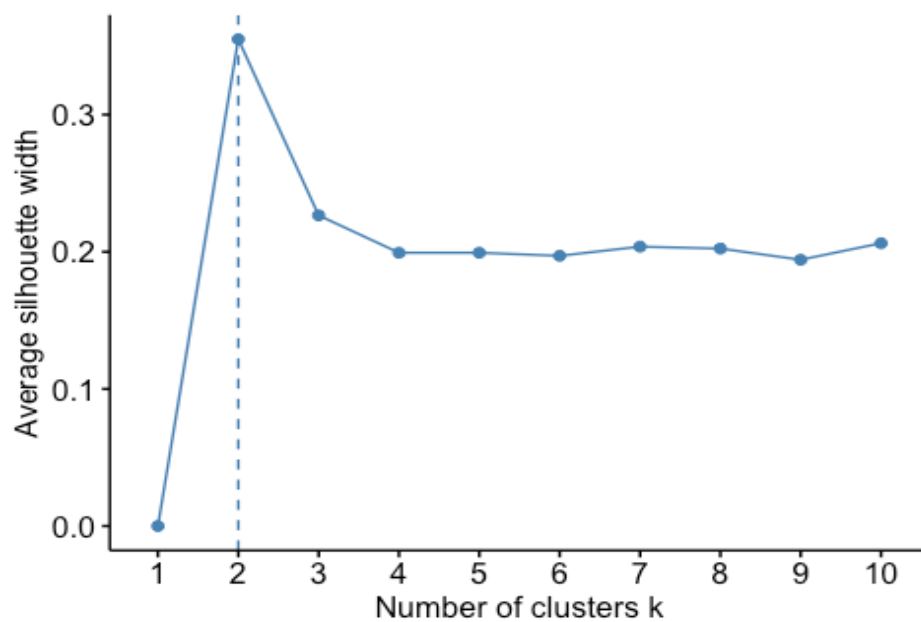**Fig.6 Optimal number of clusters-silhouette**
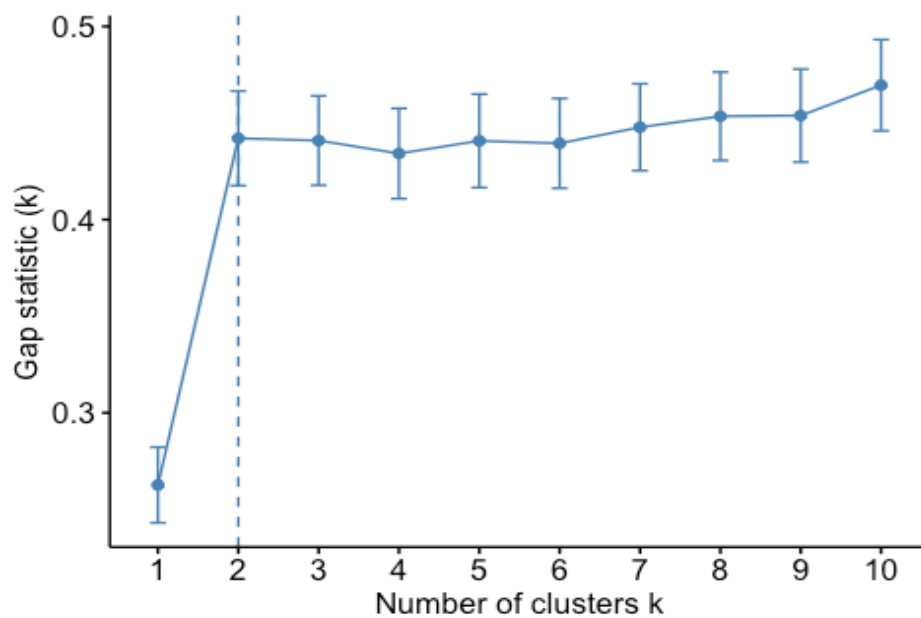

Fig.6 Optimal number of clusters-gap

Fig.7 Full display of cluster plot

Fig.8 PAM-quality

Clusters silhouette plot
Average silhouette width: 0.36