ABSTRACT

Measuring the Effect of Culture on Economic Growth and Development

Tianyu Li, M.S.
Department of Mathematical Sciences
Northern Illinois University, 2010
Alan Polansky, Director

Is there a link between economic growth and culture? The data in this paper, including data from 15 countries, suggests that there is. Culture, in this paper, was divided into seven different variables. To test if those seven variables have a significant effect on GDP growth rate, we ran multiple linear regression models by country separately and that several of the seven variables have significant effect on GDP growth rate for 12 of the countries. The analysis was also done by using mixed models where we also found that several variables affect economic growth through some factors such as labor and international trade. Also, I grouped countries into clusters, and found that countries with similar culture features often have different significant variables that affect economic growth.

NORTHERN ILLINOIS UNIVERSITY
DEKALB, ILLINOIS

AUGUST 2010

Measuring the Effect of Culture on Economic Growth and Development

BY

TIANYU LI

A THESIS SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULLFILMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICAL SCIENCES

NORTHERN ILLINOIS UNIVERSITY

Thesis Director:
Alan Polanksy

ACKNOWLEDGEMENTS

I would like to thank Prof. Polansky, who is my master's thesis advisor. He helped me to work with some ideas to analyze the data. I also would like thank Prof. Shuva Gupta, Prof. Sudhir Gupta and Prof. Peng Shi. Prof. Shuva Gupta and Prof Sudhir Gupta gave some advice on the regression part of the analysis, and Prof. Peng Shi helped me in R program of time series part. Finally, I thank my friends Peiyong Yu and Jinlang Sun who helped me with the data.

TABLE OF CONTENTS

Page

Chapter 1: Introduction

1.1 Introduction to Culture and the relationship to economic growth

Traditional growth models predict that the economic growth rate is dependent on the rate of capital and saving accumulation. These models also suggest that capital, labor and technology greatly contribute to economic development. Neither of them includes an important factor that can also affect economic growth and development, which is local culture. In a given country, institutional and political infrastructure is often the same. Many economics structures are also similar, so why is the economic growth and development in these different countries different? What variables forecast the economic future of a country? Can people's individual beliefs and morals influence their economic behavior, and thus have an effect on the economic outcome of the country? It is possible that the variations in the economy of various countries, or even regions of a single country, can be explained by the history, geography, beliefs, morals, religion, ideology and customs of its people. In fact, Thompson (2001) states, "The total complex pattern of customary human behavior, social forms and material traits embodied in thought, speech, action, and artifacts are dependent upon the human capacity for learning and transmitting knowledge, and systems of abstract thought. This will include beliefs, morals, laws, customs, opinions, religion, superstitions, and art." Growth and development are based on institutions that require a guaranteed exchange of goods, property rights and distribution of resources. These institutions have prerequisite norms or beliefs that can be accepted by most of the people that contribute to the economy of that institution. Thus culture shapes people's behavior and economic activity in a given institution. A number of studies have examined the relationship between culture and economic development. Thompson (2001) states, "The relationship of 'culture' and 'economic development' during the past fifty years can be, and has been, viewed variably as causal, correlative or relatively autonomous." Myerson (1991) and Greif (1994) suggested that "the most common meaning of culture is that it refers to the social conventions and individual beliefs that sustain Nash equilibria as focal points in repeated social interactions or when there are multiple equilibria." Acemoglu, Johnson and Robinson (2004) concluded that "culture is viewed as a key determinant of the values, preferences and beliefs of individuals and societies and, the argument goes, these differences play a key role in shaping economic performance." Tabellini (2005) found that there is causal relationship between culture and economic development. In his publication, he measured culture by indicators of individual values and beliefs.

How can culture affect economic development? It appears that it affects economic development at both the microeconomic and macroeconomic levels. Considering the former, culture affects economic development by altering people's behavior, beliefs and values. In addition, culture can affect economic development by influencing the development of civil infrastructure and political systems.

1.2 How culture affects economic growth at micro level

At the micro level, culture influences people's beliefs, values and behaviors in a way that affects economic development. Tabellini (2008) suggested that "a cultural feature often mentioned as a driver of economic development is the conviction that individual effort is likely to pay off." He explains that if people in a region are motivated and they believe that their behavior will result in certain economic benefits, the results can lead to economic success. As a result, they will take the opportunity to work hard and invest their money. By contrast, if people only believe that such things as "good luck" or similar uncertain predictors are the only things that can bring them success or fortune, they may become passive and lazy. Greif (1993) argues that, "different cultures generate different sets of beliefs about how people behave and this can alter the set of equilibria for a given specification of institutions." Thus, different cultural features contribute to the formation of different beliefs, values and behaviors, and hence different economic outcomes.

People in the East have different norms, values, beliefs and behaviors from those in West. They often view their environment and events in ways that are different from those of a Westerner. For example, during China's long history many different cultures and beliefs have evolved. The most notable ones include "Confucianism", "Taoism" and "Buddhism". Confucianism emphasizes humanity and benevolence and focuses on fostering personal cultivation and moral development. It also focuses on righteousness, politeness, wisdom and faith. Taoism teaches people to be content with one's lot and lessen people's desire for a better lot in life. Buddhism tells people to do good works and forbids the killing of animals, stealing, cheating, and drinking alcohol and sexual misconduct. In essence, all these religious beliefs attempt to stimulate the cultivation of moral character. Thus people trust and respect each other, which can have a beneficial effect on the market for the exchange of goods. However, Confucianism also emphasizes the "Golden Mean" which encourages a median approach to life, while eschewing extremes to left or right. This may confine people's initiative and creativity. Also, Confucianism has contempt for natural scientists and those who pursue handicrafts. This outlook may impede the development of the natural sciences and the handicraft industry and thus hinder the development of technology. Taoism believes that everything has its own natural pattern and so man should not break the natural order to strive for unnatural things. So, Taoists reject politics and instead choose to follow their own destiny. A similar assertion can also be found in Buddhism, in which it asserts that everything has its own destiny and that people will experience another life after death. As a result, the Buddhist should tolerate suffering in this life and prepare for the next life. This belief tends to produce a passive attitude towards to the world in general. These cultures tended to accept this destiny, which led to exploitation by feudal lords.

1.3 How culture affects economic growth at macro level

As mentioned earlier, culture may affect economic development by influencing the choice of a particular set of institutions and political systems. Confucianism historically led to the development of agricultural systems and restrained commerce in China. This restraint of commerce inhibited the development of a nation-wide market during feudal times, producing a lag in economic, scientific and technological development behind that of the Western World. An imperial examination system gave lower-ranked intellectuals an opportunity to become officials of imperial government, producing an intellectual focus on writing essays to the detriment of science and technology. Confucianism also may have contributed much to the development of China's feudal system and postponed its decay which hindered the growth of capitalism in China while at the same time; Europeans began developing capitalism during the Renaissance and Protestant Reformation.

Another example of this premise can be seen in South and North Korea after the Korean War. Originally they had a similar economic foundation and background. However, today South Korea is much more developed than North Korea. This is because the South Koreans chose a different economic path to develop their economy after the war. South Korea built a capitalistic economic institution while North Korea chose the socialist model and political systems. The reason that they made the different choice is due to the different political and economic ideology and beliefs. South Korea accepted support from United States and their ideology, and North Korea accepted support from Russia and China. North Koreans accepted Marxism and have used his theories to build their economic and political structure.

Chapter 2: Data Description

A total of fifteen countries were selected from Asia, Europe, America (North America and Latin America), Africa and Oceania, with other countries in their own continents.

Agricultural land percentage, ALP, is the percentage of land dedicated to agriculture as an indicator of whether a country is primarily agricultural or industrial and can reflect the progress of urbanization, especially for developing countries. Agricultural countries and industrial countries exhibit different growth and development patterns. Thus, the economic outcome for each country is different depending upon people's beliefs and behaviors and because urban life is different from rural life within a given country. Industrialization may be one criterion by which one can judge whether a country has reached modern standards.

Fertility rate total (FRT, births per woman) is an indication of fecundity patterns that can show the preference of a given population for birth rates and birth policies in a particular country. In China, the fertility rate in urban areas is lower than rural areas, in part, because those in rural areas, often peasants, prefer males compared to females and because males can contribute more toward farm work in the rural areas. In addition, the family name is passed on by the male heir. Such cultural features result in an imbalance in the proportion of males versus females. This data set range extending between 1961 and 2005 should indicate trends in the country's fertility rate which can be compared with economic growth and development during the same periods. This variable is also an indicator of education levels in that, a country where people are well educated has a lower fertility rate compared to countries with lower level of education. Because educated people spend more time on their education and training and beginning their career they usually have children later in life compared to people with less education. Those who have children later in life are more likely to have a smaller family compared to those who give birth early in life. In addition, a country with a prosperous and rapidly growing economy generally has lower birth rate (see the data set).

Gross domestic product rate (GDP, annual %) is the dependent variable in my model. It is a measure of economic growth and development. GDP growth annual rate not only reflects increases in economic growth, but also is an indicator of whether growth is stable and continuous. This data set range is from 1971 to 2005.

GNI (Gross National Income) GNI per capita shows the economic growth efficiency. It is an indicator of economic growth rate and the types of economic growth. It reflects the rationale behind country's economic infrastructure, its distribution of resources and use of human capital. All of these are related to popular culture.

Life expectancy at birth (LEB, total in years) predicts a countries population lifespan. The life expectancy is based on historical life expectancies for the population and economic growth conditions. It also takes into consideration the current economic conditions of the country and predicts the level of future economic growth and political policies. LEB is calculated using these data. It may also predict a population's attitude towards the future, its economic conditions and political system. For example, when a person's attitude is generally optimistic toward life and the person believes that economic conditions will improve and the political system will be stable, the individual's life expectancy is predicted to be longer. This can in turn affect people's behavior in relation to economic activities and can change the outcome of a given economy. This data set ranges from 1977 to 2005.

The Merchandise_rate is a measure of trading activity (% of GDP). The data set ranges from 1970 to 2005. The merchandise rate can be an indicator of how trade activity plays a role in a country's economic growth. Merchandise trade tells us the degree of commercialization and marketization of a country. Merchandise trade can also be a reflection of international trade cooperation in various economic fields which creates cross cultural communication. Thus local culture is affected, mixed and even changed by these types of communication. When trading with other countries, people's world outlook, ideologies and behaviors can be drastically changed. Finally, economic growth and development of a given country are affected. A country's marketization and history of market economy development depends on its political system and economic institutions; however, they also depend on the nation's cultural viewpoints and ideologies. For example, China was closed to the world before 1978; there was practically no international trade and economic cooperation and, of course, no cross cultural communication. However, after 1978 until the present day, China has become an indispensible part of the world economy. Such changes in economic policy were caused by a change in the nation's ideologies.

Variable migration measures net immigration rates. The data set contains 45 years of net migration numbers starting with 1961 till 2005. A high immigration rate in the country not only suggests that a country's economy is prosperous, has advanced economic institutions and loose immigration policies, but also indicates that the country has cultural components that are attractive to other countries. So a country with big consecutive net immigration history has diversification as its culture features and is also rich in human resource. Usually, a country that attracts many immigrants has three important characteristics. First, it usually has a good educational system, especially at the college level. A second feature is a relative good economic environment. Finally, such countries usually have a long immigration history. A good example of this concept is the USA. It satisfies all three conditions above: good colleges, an advanced economic system and an immigration culture. Because of immigration, the USA has competitive advantages in the world market in spite of some negative factors brought by immigrants.

The last variable in the model is population growth (PopGrowth, annual %). This data set ranges from 1961 to 2005. The annual population growth rate is calculated as birth rate minus death rate. We use this to predict the future populations and an indicator of a populace's age stratification. For example, a consecutive negative population growth annual rate implies the populace's age stratification will enter into a gerontological period. Since the old and young have different culture backgrounds and belief systems, their moral behavior and world outlook are also different. Hence those differences have influence on economic growth and development and may in turn affect a given economic outcome. The stability of population annual growth rate can also reflect the stability of economic growth and the political system in a country.

Chapter 3: Methodologies

Because culture is not a quantitative variable, and it also cannot be quantified, I have used instrumental variables to substitute as culture in this paper. Culture has many features and can be divided into many different components. I have selected seven variables correlating to culture in my regression models. They all reflect some features of culture. Because of data limitations and budget, I was not able to collect all the variables correlating to culture and create survey questionnaires.

3.1 Cluster

I have also applied cluster analysis in this paper. I have grouped observations into different clusters by their patterns and characteristics. The goal is to find optimal groups that countries have within each cluster that are similar in culture features, while betweens clusters countries have culture features that are dissimilar. I have then compared the regression results by country within each cluster to see if the variables that are significant on economic growth in one country are still significant in the other countries in the cluster. Also, I will compare the different clusters to see if different culture features can cause different economic growth. In this part, I used hierarchical method.

1. Single linkage method. In this method, the distance between two clusters A and B is defined as the minimum distance between a point in A and a point in B:

   $$D(A, B) = \min\{d(\mathbf{y}_i, \mathbf{y}_j), \text{ for } \mathbf{y}_i \text{ in } A \text{ and } \mathbf{y}_j \text{ in } B\}$$

2. Average linkage approach. The distance between two clusters A and B is defined as the average of the $n_A n_B$ distances between the n_A points in A and the n_B points in B:

$$D(A, B) = \frac{1}{n_A n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(\mathbf{y}_i, \mathbf{y}_j),$$

3. Ward's method. Also called the incremental sum of squares method, it uses the within-cluster (squared) distances and the between-cluster (squared) distances:

$$\mathrm{SSE}_A = \sum_{i=1}^{n_A} (\mathbf{y}_i - \overline{\mathbf{y}}_A)'(\mathbf{y}_i - \overline{\mathbf{y}}_A), \quad \mathrm{SSE}_B = \sum_{i=1}^{n_B} (\mathbf{y}_i - \overline{\mathbf{y}}_B)'(\mathbf{y}_i - \overline{\mathbf{y}}_B),$$

$$\mathrm{SSE}_{AB} = \sum_{i=1}^{n_{AB}} (\mathbf{y}_i - \overline{\mathbf{y}}_{AB})'(\mathbf{y}_i - \overline{\mathbf{y}}_{AB}),$$

where $\overline{\mathbf{y}}_{AB} = (n_A \overline{\mathbf{y}}_A + n_B \overline{\mathbf{y}}_B)/(n_A + n_B)$

The definitions and formulas above are from Rencher[1] (2002).

3.2 Multiple Linear Regression and Time Series

For the multiple linear regression models part, I ran the regression by each country separately to find whether those variables have significant effect on GDP growth rate in order to decide whether culture can cause economic growth. I ran the multiple linear regressions to select an appropriate model by using stepwise method. After that, I plotted residuals to year, dependent variable and each independent variable separately to see if there is any violation in constant variance and independence. Then I plotted dependent variable to each independent variable to see if there is linear relationship between them, and also if there is curve in the graphs to indicate some quadratic forms should be included in my regression models. If some graphs showed curves between dependent variable and independent variables, I will include the corresponding variable with its squared form. After adding the quadratic form to my regression model, I did the model selection by stepwise again, and meanwhile, I also used CP method to select model to compare the results and then determine which variables should be included in my model. Finally, I ran the model that I

[1] "Methods Of Multivariate Analysis" -John Wiley & Sons (2 Ed) (Wiley Series In Probability And Statistics)

selected in the last step to see which variables are significantly effective in my regression. In the end of this part, I ran R program to check autocorrelations and time lags of each single variable for each country.

3.1 Mixed Models

Following these analyses, I used mixed model which included all countries to test if country and year have significant effects on economic growth. First, I ran the fixed models. I used stepwise method to select model including all countries and determined if I should include any interaction variable in my fixed model. Then I ran the fixed models with interaction terms versus without interaction terms. After that I also added "repeated year" to the fixed models, and last I compared AIC and BIC values of those four models to determine the fixed model in order to test if any explanatory variables have a significant effect on economic growth. In mixed models part, I also have four models with different variance-covariance matrices, which are "Compound Symmetry," "Unconstructed" and "Heterogeneous Compound Symmetry," and the last one with "random countries" as an option in my SAS program codes.

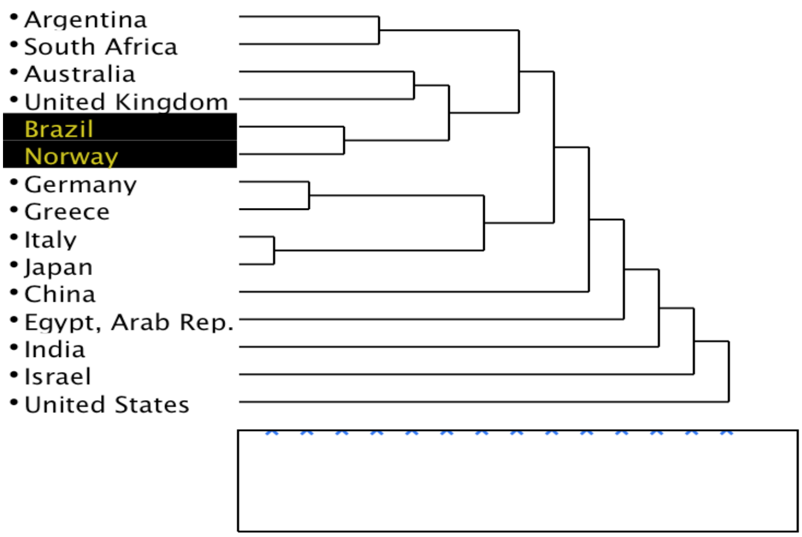
Since there are some missing values in my data set, I used "proc mixed" program here because "proc mixed" can fix missing observations by within and between methods. For example, the variable FRT began in 1961 but the variable LEB only began in 1977.

Chapter 4: Result

In this part, I will show the results of process of clusters, multiple regressions and mixed models. And I will also give the explanation corresponding to the different results and output respectively.

4.1 The Results of Clusters

4.1.1 Results of Cluster by Single Linkage

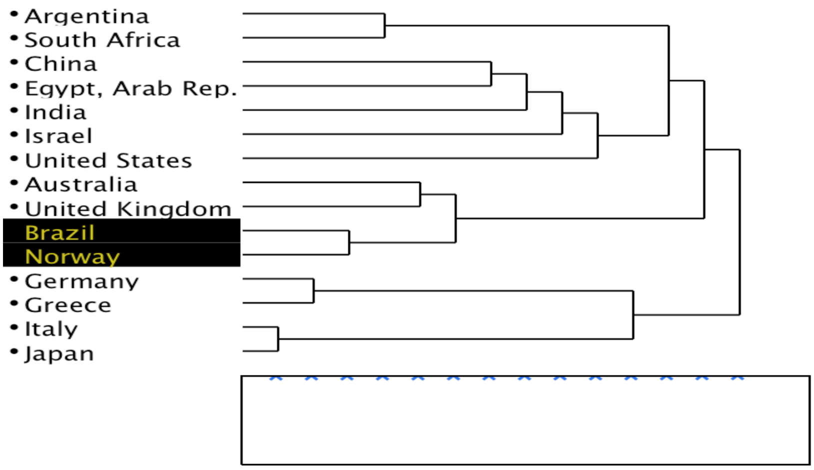


Clustering History

Number of Clusters	Distance	Leader	Joiner
14	22.04540769	Italy	Japan
13	22.06807649	Germany	Greece
12	22.06807649	Australia	Italy
11	22.06807649	Australia	Germany
10	22.06807649	Brazil	Norway
9	22.06807649	Argentina	South Africa

8	22.06807649	Argentina	Brazil
7	22.06807649	Australia	United Kingdom
6	22.06807649	Argentina	Australia
5	22.09072203	Argentina	China
4	22.09072203	Argentina	Egypt, Arab Rep.
3	22.09072203	Argentina	India
2	22.09072203	Argentina	Israel
1	22.09072203	Argentina	United States

4.1.2 Results of Cluster by Average

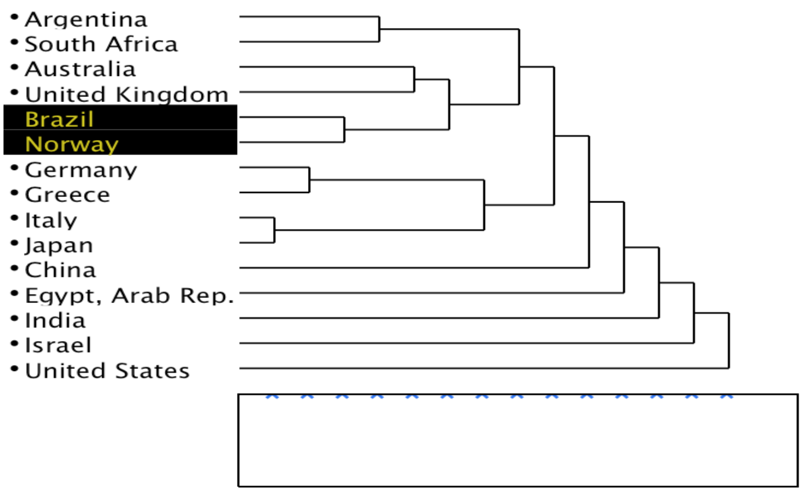


Clustering History

Number of Clusters	Distance	Leader	Joiner
14	22.04540769	Italy	Japan
13	22.06807649	Germany	Greece
12	22.06807649	Brazil	Norway
11	22.06807649	Argentina	South Africa
10	22.06807649	Australia	United Kingdom
9	22.07940217	Australia	Brazil

8	22.07940217	Germany	Italy
7	22.08789261	Argentina	Australia
6	22.08977893	Argentina	Germany
5	22.09072203	Argentina	China
4	22.09072203	Argentina	Egypt, Arab Rep.
3	22.09072203	Argentina	India
2	22.09072203	Argentina	Israel
1	22.09072203	Argentina	United States

4.1.3 Results of Cluster by Wards



Clustering History

Number of Clusters	Distance	Leader	Joiner
14	15.58845727	Italy	Japan
13	15.60448653	Germany	Greece
12	15.60448653	Brazil	Norway
11	15.60448653	Argentina	South Africa
10	15.60448653	Australia	United Kingdom
9	15.62049935	Australia	Brazil

8	15.62049935	China	Egypt, Arab Rep.
7	15.62049935	China	India
6	15.62049935	China	Israel
5	15.62049935	China	United States
4	15.62849961	Germany	Italy
3	15.63192704	Argentina	China
2	15.63961005	Argentina	Australia
1	15.65233064	Argentina	Germany

The smallest distance is between Japan and Italy in all of those three graphs above. Most of the clusters remain in the same groups. By the results of multiple linear regression models in the next section, I found that the countries with similar cultural features include different significant variables that can affect their economic growth rates.

4.2 Multiple Linear Regression Models

In this process I firstly generated five interaction variables among the variables as followed: X_1 is the interaction of FRT and PopGrowth, X_2 is the interaction of PopGrowth and Migration, X_3 is the interaction of GNIpercap and Merchandise_rate, X_4 is the interaction of PopGrowth and LEB, and X_5 is the interaction of ALP and GNIpercap. Then the full model (without quadratic form) is as follow:

$$GDPrate = \beta_0 + \beta_1 ALP + \beta_2 FRT + \beta_3 GNIpercap + \beta_4 LEB + \beta_5 Merchandise\_rate \\ + \beta_6 Migration + \beta_7 PopGrowth + \beta_8 X_1 + \beta_9 X_2 + \beta_{10} X_3 + \beta_{11} X_4 + \beta_{12} X_5 + \varepsilon$$

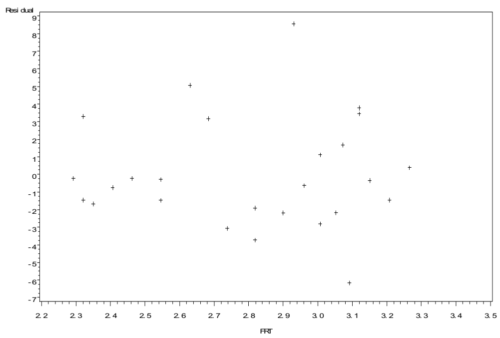
4.2.1 Argentina

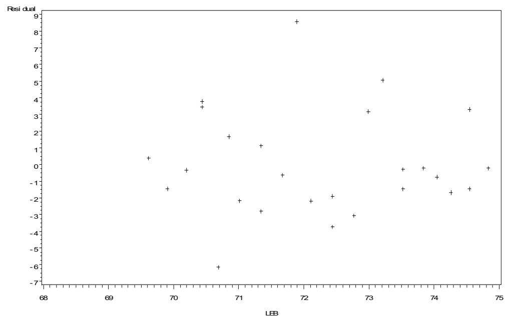
The result of stepwise and the normal plots

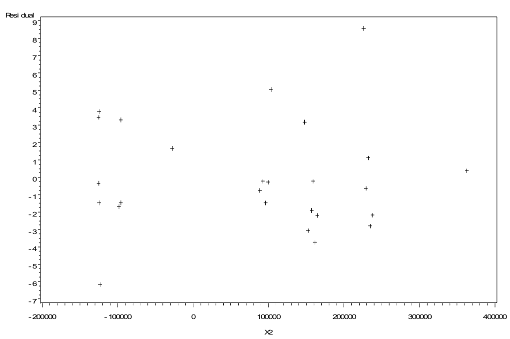
Step	Variable Entered	Variable Removed	Label	Number Vars in	Partial R-Square	C(p)	F Value
1	GNIpercap		GNIpercap	1	0.1026	24.5085	2.75
2	X1		2	0.1609	0.2635	18.1713	5.02

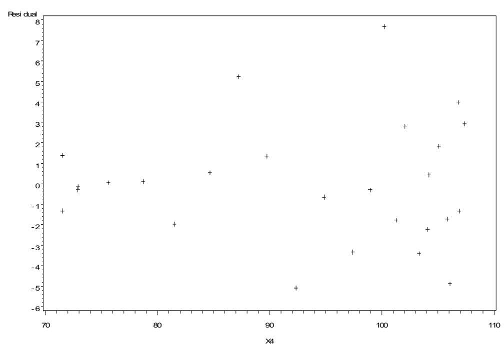
- 9+ * ++
- | ++++++
- | *+*+*++
- | +***++
- | ++***+*
- | *+******
- | *+***+
- -5+ * ++*+*+
- +----+----+----+----+----+----+----+----+----+
- -2 -1 0 +1 +2

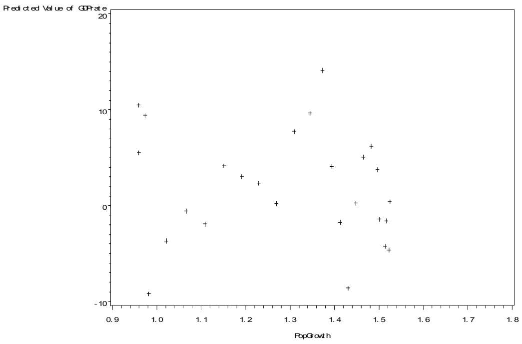
Some plots of residuals versus variables and dependent variable versus independent variables:

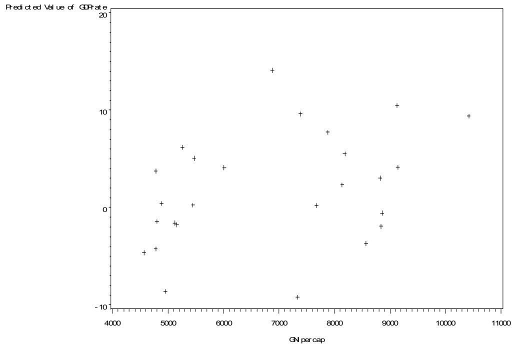












Seeing the first part graphs of plots, I can conclude that there is no violation of constant variance and independence. Thus the assumptions all hold. Based on the graphs of plots of second part, I found that there seems to be a curve between GDP growth rate and PopGrowth, and between GDP growth rate and GNIpercap. So I decide to add two variables' squared term into my regression model. Then I used CP method to decide how many variables and which variables I should include in my model, and it suggests to include six variables: FRT, Merchandise_rate, X3, X4, X5 and X7 with C(p)= 6.9846 and R-Square=0.6531. Finally, I ran the regression model including those six variables and found the result as below:

Root MSE	4.44878	R-Square	0.6531
Dependent Mean	1.88089	Adj R-Sq	0.5436
Coeff Var	236.52571		

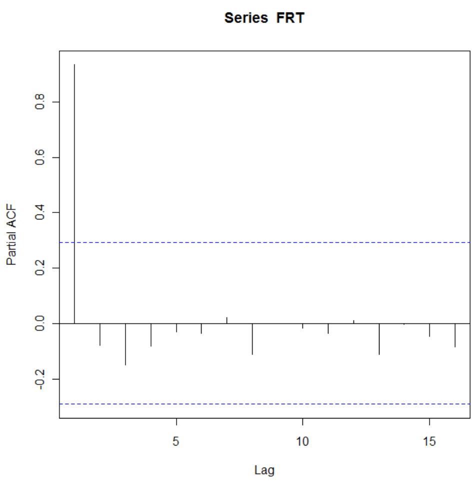
Durbin-Watson D	2.030
Number of Observations	26
1st Order Autocorrelation	-0.025

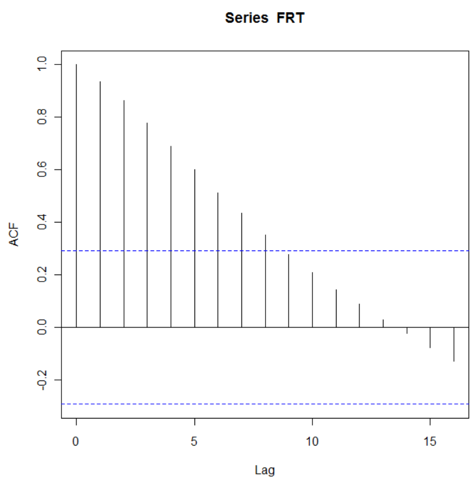
Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-244.25035	55.51237	-4.40	0.0003
FRT*	FRT	1	84.14521	29.87473	2.82	0.0110
Merchandise_rate	Merchandise rate	1	-3.19426	0.88989	-3.59	0.0020
X3		1	0.00044578	0.00010603	4.20	0.0005
X4		1	-0.98291	0.70272	-1.40	0.1780
X5		1	0.00068936	0.00017609	3.91	0.0009
GNIpercap^2		1	-0.00000240	6.157327E-7	-3.90	0.0010

*The variables in red are significant in this paper.

The graphs of time lags for FRT by ACF and PACF:





So there is time lag in ACF for FRT and the lag is 8.

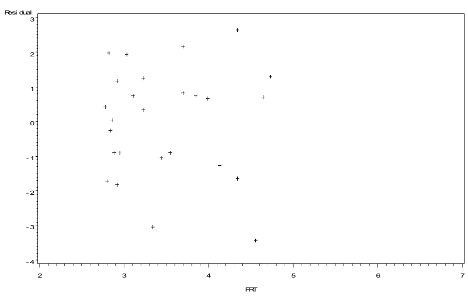
4.2.2 South Africa

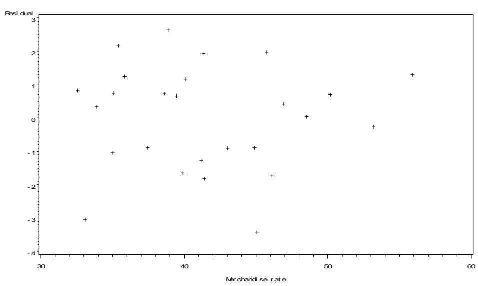
The result of stepwise and the normal plot:

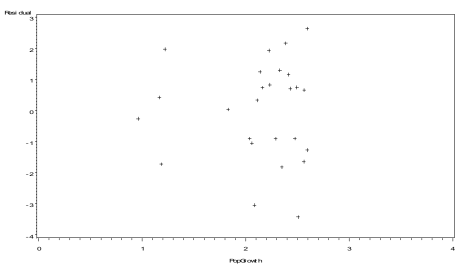
South AfricaStep	Variable Entered	Variable Removed	Label	Number Vars in	Partial R-Square	C(p)	F Value
1	X3			1	0.2928	-2.9922	9.94

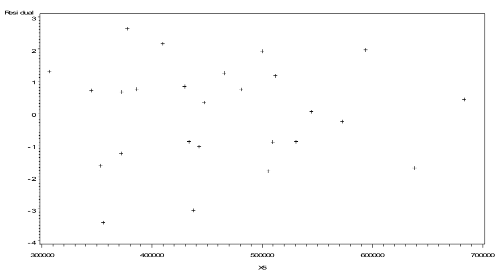
```
2.5+                                  ++*++++*
   |                            ***+*+*
   |                      ******+*
-0.5+                 *+***+
   |          *+*+***
   |     +++++++
-3.5++++++*     *
    +----+----+----+----+----+----+----+----+----+----+
        -2        -1        0        +1        +2
```

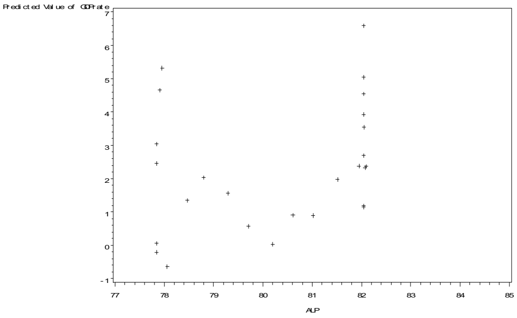
Some plots of residuals versus variables and dependent variable versus independent variables:

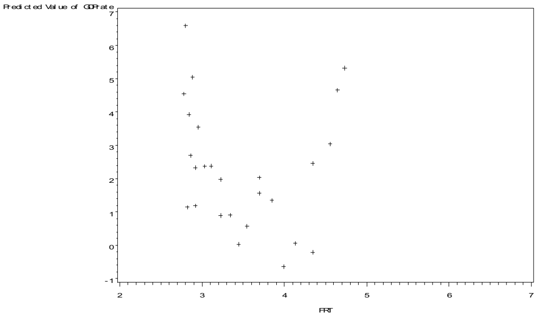












By the result of the plots above, the assumption of constant variance holds. And also, I found that there seems to be a curve relation between predicted GDPrate and ALP, and predicted GDPrate and FRT. So I include the two variables' squared term in my regression models. And then I use C(p) and stepwise to determine a model. C(p) suggests to include 12 variables: ALP, FRT, LEB, Merchandise_rate, PopGrowth, GNIpercap, X1, X2, X3, X5, X6 and X7 with R-Squared=0.4398 and C(p)=13.1299. Then I ran the multiple linear regressions with above variables and found none of the variables above are significant. And stepwise still only included one variable, X3, left in the regression model. So I ran the regression model with a single variable X3. Then the SAS output listed below:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	46.31171	46.31171	11.39	0.0019
Error	33	134.17387	4.06587		
Corrected Total	34	180.48558			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-6.05099	2.55289	-2.37	0.0238
X3		1	0.00255	0.00075636	3.37	0.0019

Durbin-Watson D	1.724
Number of Observations	35
1st Order Autocorrelation	0.113

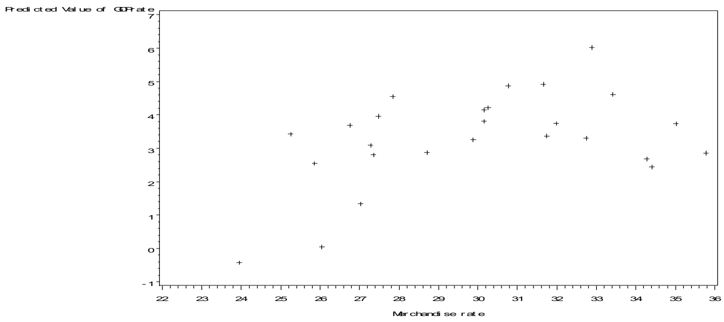
4.2.3 Australia

The result of stepwise and the normal plot:

```
 2.75+                                    * ++++
     |                                 *  ++++
     |                               *++++
     |                             ++*++
     |                          +****
 0.25+                        +****
     |                     ++++*
     |                 ****+****
     |              *+*++
     |     *   ++*+
-2.25+     ++++
      +----+----+----+----+----+----+----+----+----+----+
```

-2 -1 0 +1 +2

Some plots of residuals versus variables and dependent variable versus independent variables:



By the result of the plots above, the assumption of constant variance holds. And also, I found that there seems to be a curve relation between predicted GDPrate and Merchandise_rate, so I include the one variables' squared term in my regression models. And then I use C(p) and stepwise to determine a model. C(p) suggested to determine the model with variables as ALP, LEB, Merchandise_rate, PopGrowth, GNIpercap, X1, X2, X3, X5, X6. Then I ran the regression with the variables above and obtain the output as followed:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	57.59021	5.75902	3.57	0.0133
Error	15	24.21961	1.61464		
Corrected Total	25	81.80982			

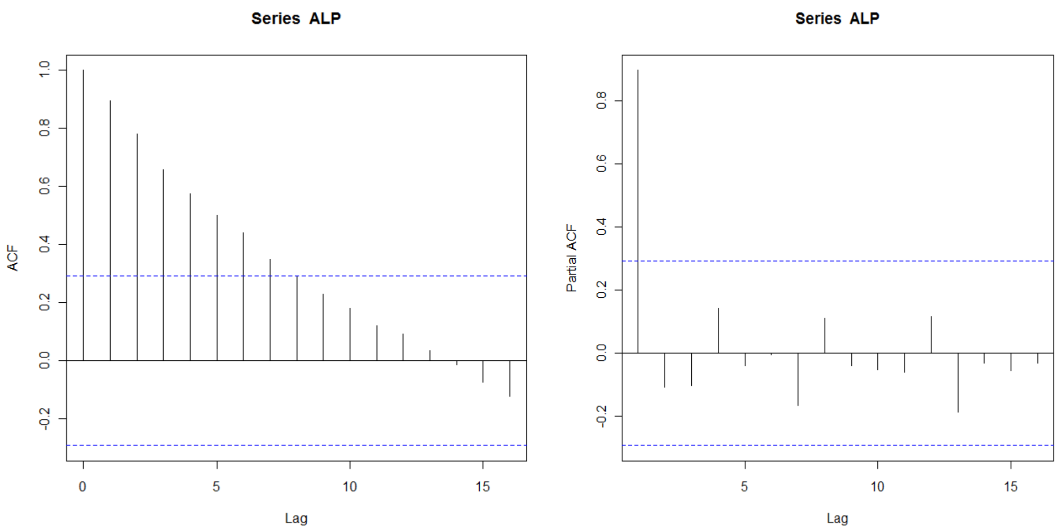
Durbin-Watson D	3.232
Number of Observations	26
1st Order Autocorrelation	-0.627

Parameter Estimates

Variable	Label	DF	Parameter	Standard	t Value	Pr > |t|

			Estimate	Error		
Intercept	Intercept	1	354.41240	198.27732	1.79	0.0941
ALP	ALP	1	-4.75466	1.63967	-2.90	0.0110
LEB	LEB	1	-3.84628	1.86849	-2.06	0.0573
Merchandise_rate	Merchandise rate	1	18.19388	4.65196	3.91	0.0014
PopGrowth	PopGrowth	1	-25.52072	17.81438	-1.43	0.1725
GNIpercap	GNIpercap	1	-0.01747	0.00572	-3.05	0.0081
X1		1	11.18943	8.79124	1.27	0.2225
X2		1	0.00000455	0.00000304	1.50	0.1552
X3		1	0.00020312	0.00007822	2.60	0.0202
X5		1	0.00020164	0.00006405	3.15	0.0066
Merchandise_rate^2		1	-0.35689	0.09920	-3.60	0.0026

The graphs of time lags for ALP by ACF and PACF:



There exists time lag by the graph above and the lag is 7 for ALP.

4.2.4 U.K.

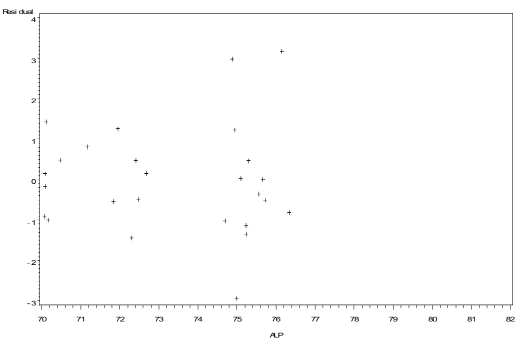
The result of stepwise and the normal plot:

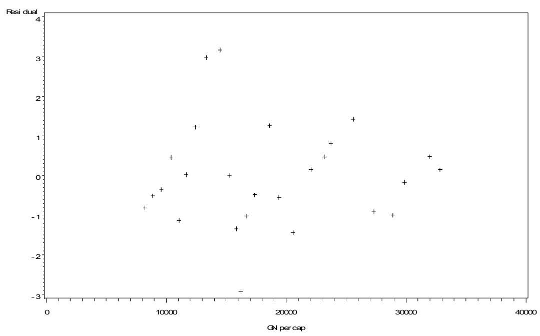
Summary of Stepwise Selection

Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	FRT		FRT	1	0.1489	0.1489	-1.0466	4.20	0.0515

```
  3.5+                                    *  +++
     |                                *+++++++
     |                           +*+*+*++
  0.5+                       *****+***
     |                  ***+*****
     |              *+*+*+*+
 -2.5+      ++*++++
      +----+----+----+----+----+----+----+----+----+----+
          -2       -1        0       +1       +2
```

Some plots of residuals versus variables and dependent variable versus independent variables:





By the result of the plots above, the assumption of constant variance does not holds with variables ALP and GNIpercap. I will determine my model first to see if those above four variables will appear in my model, then I will deal with those they do. And also, I found that there seems to be no curve relation between predicted GDPrate and explanatory variables, so I did not include any variables' squared term in my regression models. And then I use C(p) to determine a model. Unfortunately, the models I selected by C(p) did not include any significant variable. Then I ran the regression with FRT and obtain the output as follows:

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	2.99466	4.09416	0.73	0.4697
FRT	FRT	1	-0.34596	2.26046	-0.15	0.8793

There is no significant variable for UK by regression result.

4.2.5 Brazil

The result of stepwise and the normal plot:

- No variable met the 0.1500 significance level for entry into the model.

```
 3.5+                                   *  +++
    |                              *+++++++
    |                        ++*+*+*++
 0.5+                  *****+***
    |           ***+*****
    |     *+*+*+*+
-2.5+ ++*++++
     +----+----+----+----+----+----+----+----+----+----+
         -2        -1         0        +1        +2
```

By the result of the plots, the assumption of constant variance holds. And also, I found that there seems to be no curve relation between predicted GDPrate and explanatory variables, so I did not include any variables' squared term in my regression models. And then I used C(p) to determine a model. Then C(p) showed that there are three models whose C(p) values are very close to K+1; then I ran the three regression models and selected one of them as my regression model for Brazil. The graphs of time lags for the significant variables:

Analysis of Variance

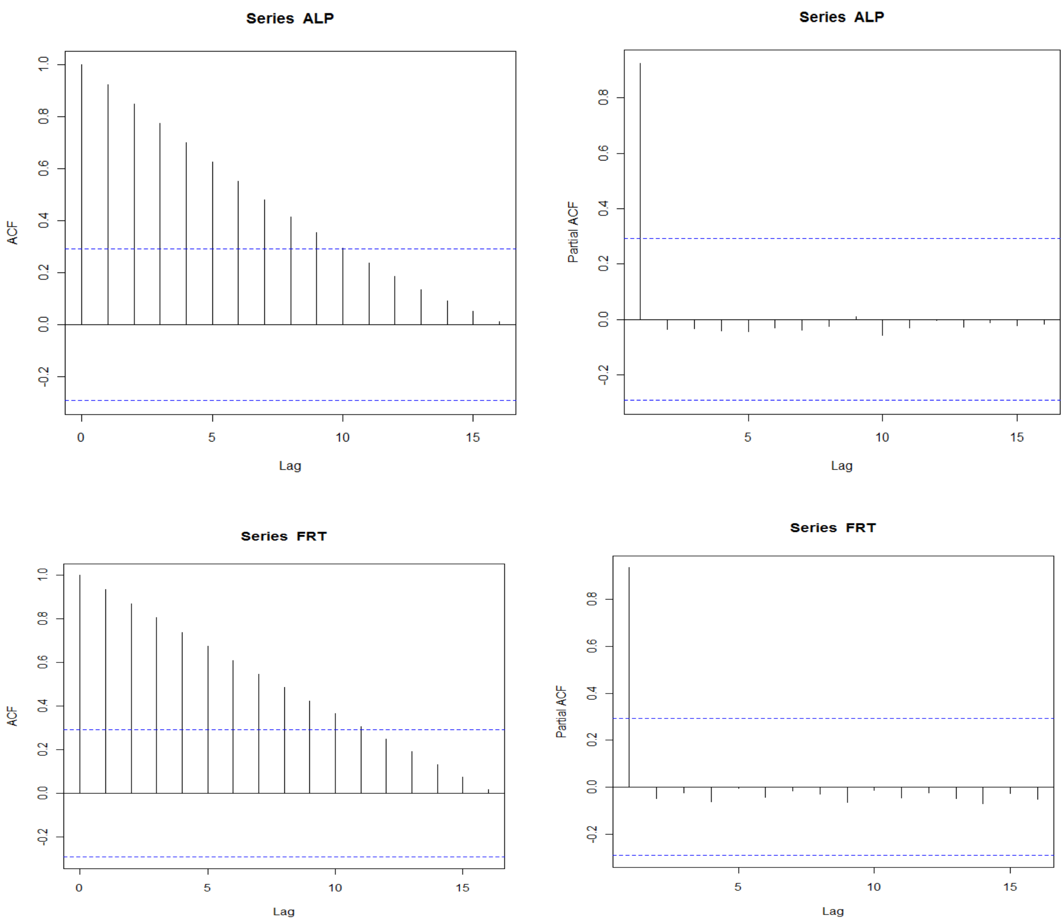
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	96.39203	13.77029	1.84	0.1487
Error	16	119.99069	7.49942		
Corrected Total	23	216.38273			

Durbin-Watson D	1.205
Number of Observations	24
1st Order Autocorrelation	0.378

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-791.62434	309.81493	-2.56	0.0212
ALP	ALP	1	24.38449	9.95986	2.45	0.0263
FRT	FRT	1	23.32956	9.19537	2.54	0.0220
Migration	Migration	1	-0.00011939	0.00009708	-1.23	0.2366
GNIpercap	GNIpercap	1	0.11451	0.03933	2.91	0.0102
X2		1	0.00010288	0.00006800	1.51	0.1498
X3		1	0.00011371	0.00007447	1.53	0.1463
X5		1	-0.00381	0.00136	-2.80	0.0129

The graphs of time lags for ALP and FRT by ACF and PACF:



Time lag exists for both ALP and FRT, and the lags of ALP are 9 and the lags of FRT are 10.

4.2.6 Norway

The result of stepwise and the normal plot:

Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)
1	Merchandise_rate		Merchandise rate	1	0.1596	0.1596	3.1673

```
   2.25+                                   +*++
       |                               +++++
       |                            +*+* *
       |                       **+****
       |                     ***++
 -0.25+                   * ****
       |               +**++
       |           *+*+*
       |        +++*+
       |    ++++
 -2.75++++  *
      +----+----+----+----+----+----+----+----+----+----+
          -2        -1         0        +1        +2
```

By the result of the plots above, the assumption of constant variance holds. And also, I found that there seems to be no curve relation between predicted GDPrate and explanatory variables, so I did not include any variables' squared term in my regression models. And then I use C(p) to determine a model. Then I ran the linear regression model including the variables selected by C(p) method.

Number in Model	R-Square	C(p)	Variables in Model
4	0.2984	5.0114	FRT LEB GNIpercap X3

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	20.98109	5.24527	2.23	0.1001
Error	21	49.33447	2.34926		
Corrected Total	25	70.31556			

Durbin-Watson D	1.169
Number of Observations	26
1st Order Autocorrelation	0.408

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-202.94692	92.31504	-2.20	0.0393
FRT	FRT	1	-2.39523	4.15752	-0.58	0.5707
LEB	LEB	1	2.82206	1.25635	2.25	0.0356
GNIpercap	GNIpercap	1	-0.00093565	0.00032501	-2.88	0.0090
X3		1	0.00001136	0.00000494	2.30	0.0320

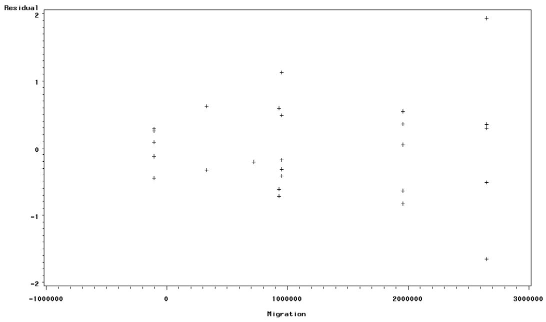
4.2.7 Germany

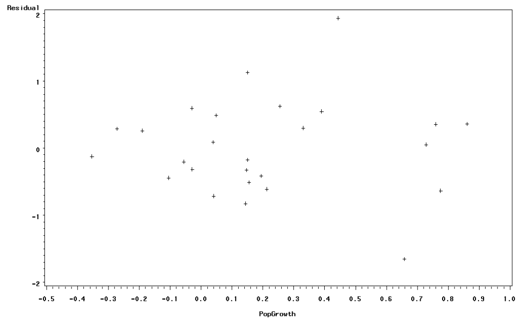
The result of stepwise and the normal plot:

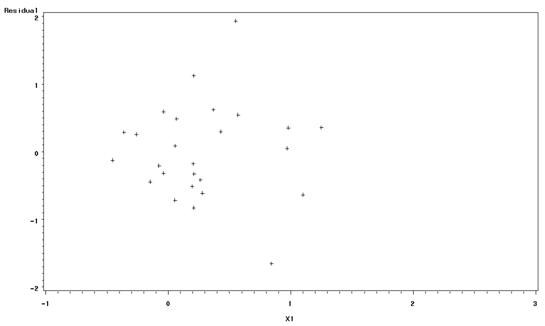
- No variable met the 0.1500 significance level for entry into the model.

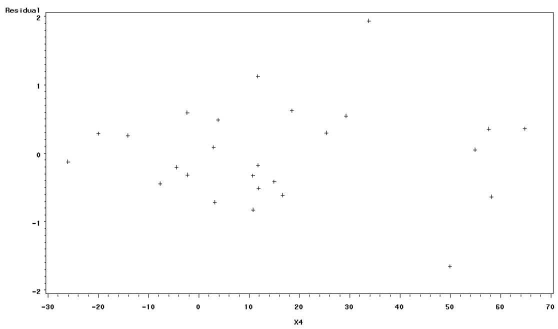
```
Normal Probability Plot
             1.75+                          * ++++
                 |                       *++++++
                 |                    +*+*+*+
                 |                 *****+***
                 |              **+*****
                 |           *+*+*+**
                 |    +++++++
            -1.75+++++ *
                  +----+----+----+----+----+----+----+----+----+----+
                      -2        -1         0        +1        +2
```

Some plots of residuals versus variables and dependent variable versus independent variables:









By the result of the plots above, the assumption of constant variance does not hold. The variables Migration, PopGrowth, X1 and X4 violate the assumption. And also, I found that there seems to be no curve relation between predicted GDPrate and explanatory variables, so I did not include any variables' squared term in my regression models. And then I used C(p) to determine a model. Then I ran the linear regression model including the variables selected by C(p) method.

Number in Model	R-Square	C(p)	Variables in Model
6	0.6934	6.8627	ALP Merchandise_rate GNIpercap X1 X4 X5

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	30.76777	7.69194	5.77	0.0027
Error	21	27.98488	1.33261		
Corrected Total	25	58.75266			

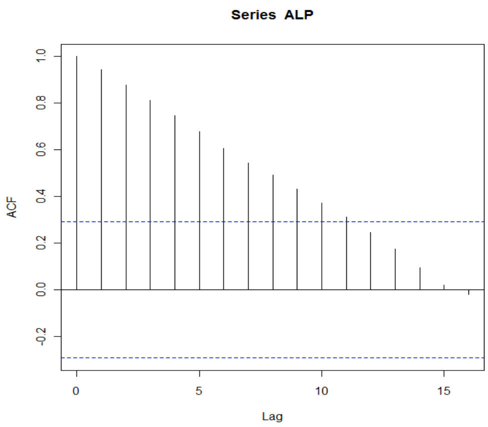
Durbin-Watson D	1.644
Number of Observations	26
1st Order Autocorrelation	0.088

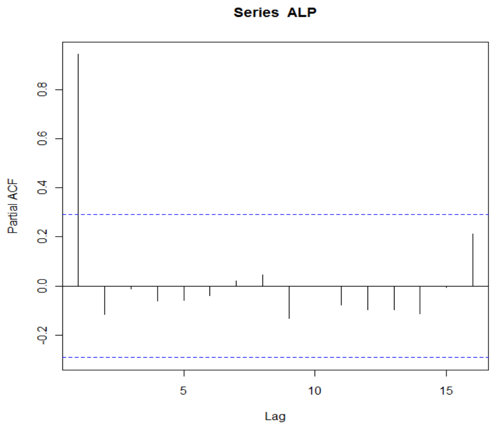
Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	217.49779	55.11477	3.95	0.0007
ALP	ALP	1	-4.34457	1.10234	-3.94	0.0007
Merchandise_rate	Merchandise rate	1	0.16422	0.06519	2.52	0.0199

GNIpercap	GNIpercap	1	-0.01170	0.00257	-4.56	0.0002
X5		1	0.00022983	0.00005023	4.58	0.0002

The graphs of time lags for ALP and FRT by ACF and PACF:



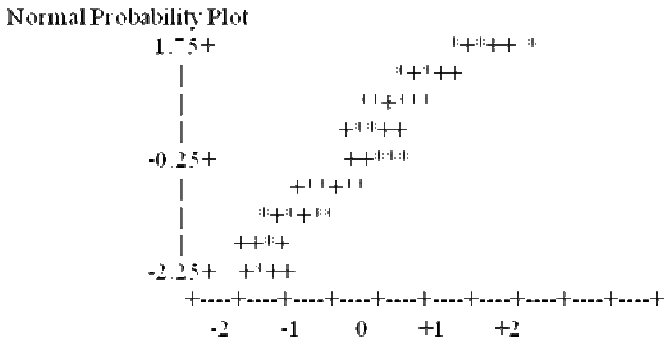


The time lag exists for AlP and the lags are 10.

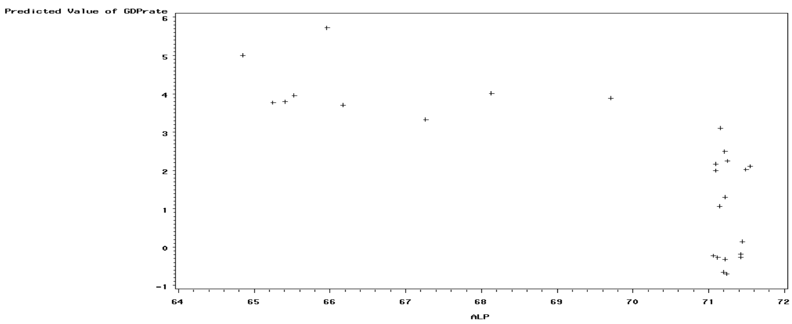
4.2.8 Greece

The result of stepwise and the normal plot:

Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)
1	X3			1	0.5008	0.5008	2.6793



Some plots of residuals versus variables and dependent variable versus independent variables:



By the result of the plots above, the assumption of constant variance holds. And also, I found that there is a curve relation between predicted GDPrate and ALP, so I did include this variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. But C(p) did not give me any regression model including any significant variable. Thus I ran the regression model only including X3.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	62.31911	62.31911	24.08	<.0001

Error		24	62.11367	2.58807		
Corrected Total		25	124.43277			
Durbin-Watson D			2.062			
Number of Observations			26			
1st Order Autocorrelation			-0.044			
Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-2.60622	0.99904	-2.61	0.0154
X3		1	0.00001056	0.00000215	4.91	<.0001

4.2.9 Italy

The result of stepwise and the normal plot:

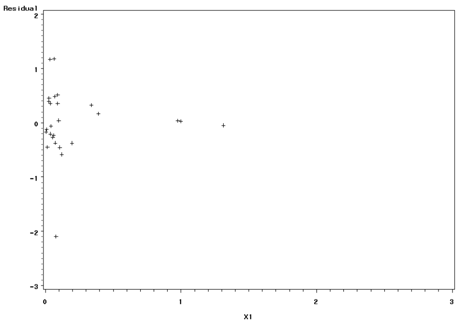
Summary of Stepwise Selection

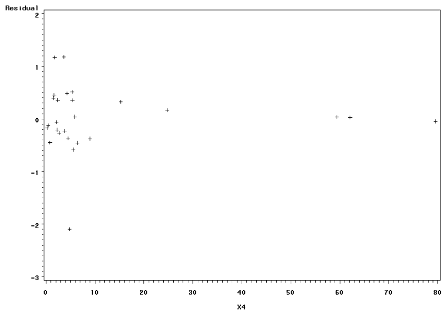
Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	ALP		ALP	1	0.1338	0.1338	22.3497	3.71	0.0661

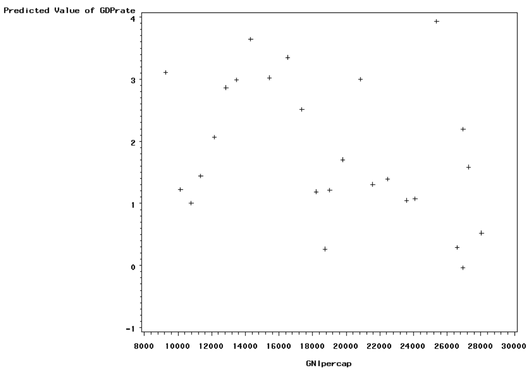
2	FRT	FRT	2	0.1367	0.2705	17.3493	4.31	0.0492
3	Merchandise _rate	Merchandis e rate	3	0.1021	0.3726	14.1208	3.58	0.0717

```
Normal Probability Plot
          1.25+                           * +++*++++
              |                       ++++*+++
              |                    *****+**** *
              |               * * ****+*****
              |          +*++++++
              | ++++++++
              |+
         -2.25+     *
               +----+----+----+----+----+----+----+----+----+----+
                   -2       -1        0       +1       +2
```

Some plots of residuals versus variables and dependent variable versus independent variables:







By the result of the plots above, the assumption of constant variance does not holds. The variables X1 and X4 violate the assumption. I found that there is a curve relation between predicted GDPrate and GNIpercap, so I did include this variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. But C(p) did not give me any regression model including any significant variable. Thus I ran the regression model only including ALP, FRT and Merchandise_rate.

The graphs of time lags for ALP and FRT by ACF and PACF:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	28.62664	9.54221	2.68	0.0638
Error	31	110.21425	3.55530		
Corrected Total	34	138.84089			

Durbin-Watson D	2.127
Number of Observations	35
1st Order Autocorrelation	-0.079

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-11.75433	8.01012	-1.47	0.1523
ALP	ALP	1	0.19693	0.14193	1.39	0.1752

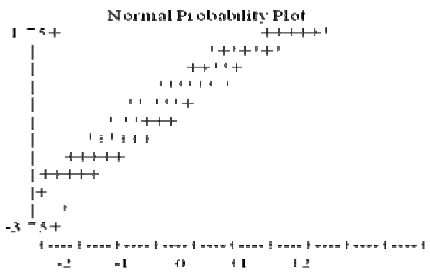
FRT	FRT	1	0.96291	1.16608	0.83	0.4152
Merchandise_rate	Merchandise rate	1	0.04345	0.07373	0.59	0.5599

4.2.10 Japan

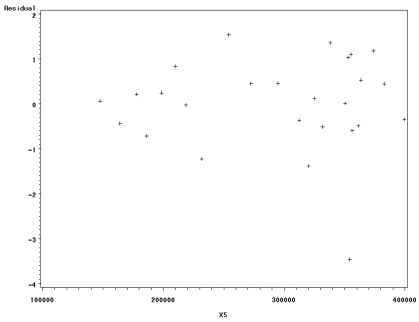
The result of stepwise and the normal plot:

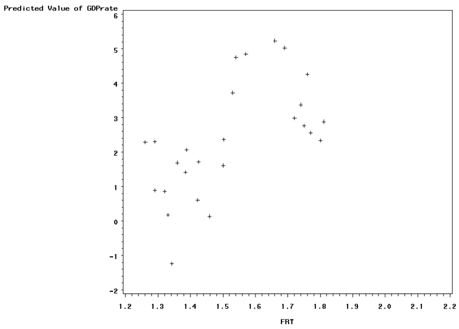
Summary of Stepwise Selection

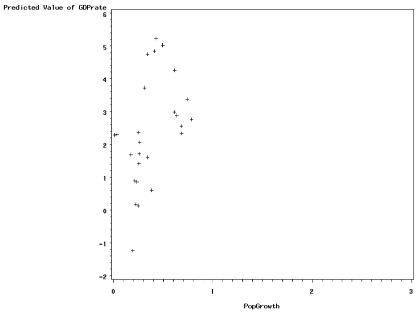
Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Migration		Migration	1	0.2746	0.2746	8.3343	9.08	0.0060
2	FRT		FRT	2	0.2472	0.5217	-0.0015	11.89	0.0022

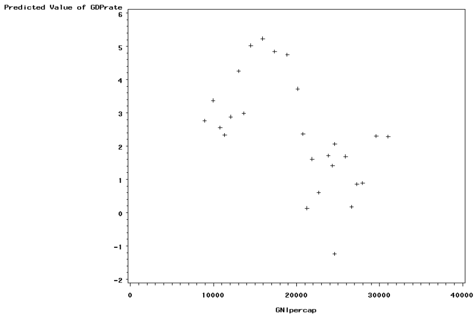


Some plots of residuals versus variables and dependent variable versus independent variables:









By the result of the plots above, the assumption of constant variance does not hold. The variables X5 violate the assumption. I found that there are curve relations between predicted GDPrate and FRT, GDPrate and PopGrowth, GDPrate and GNIpercap, so I did include those variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. C(p) suggested to include variables ALP and X2. I also added FRT and Migration as suggested by stepwise, but migration is insignificant so I omitted it and only added FRT. Thus I ran the regression model only including ALP, FRT and X2.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	103.41032	34.47011	11.80	<.0001

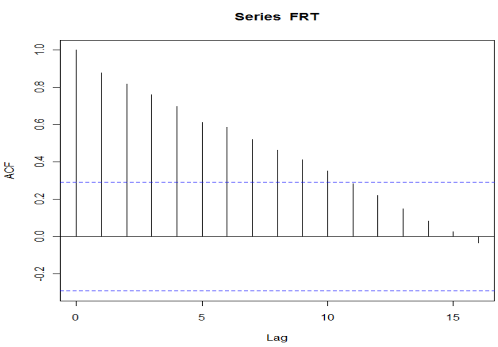
Error	31	90.54872	2.92093
Corrected Total	34	193.95904	

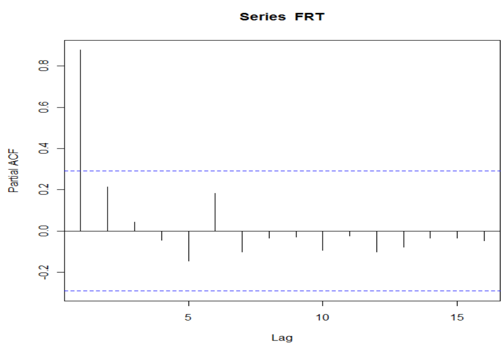
Durbin-Watson D	1.633
Number of Observations	35
1st Order Autocorrelation	0.163

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-5.07787	4.74294	-1.07	0.2926
ALP	ALP	1	-0.87464	0.66329	-1.32	0.1969
FRT	FRT	1	13.71265	4.14392	3.31	0.0024
X2		1	-0.00000520	0.00000137	-3.79	0.0007

The graphs of time lags for ALP and FRT by ACF and PACF:



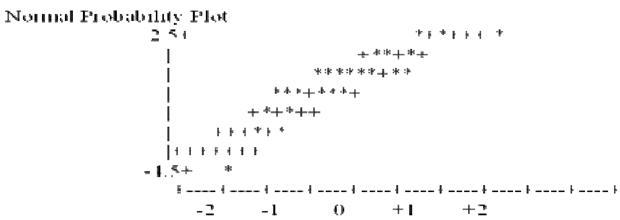


The graphs show that time lag exists for FRT and the lags are 10.

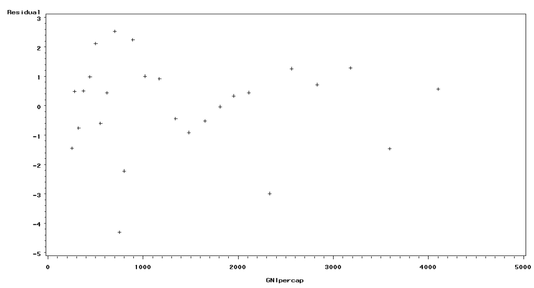
4.2.11 China

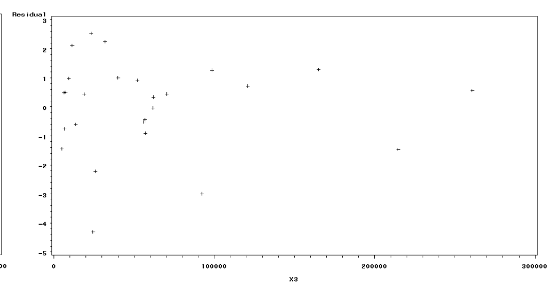
The result of stepwise and the normal plot:

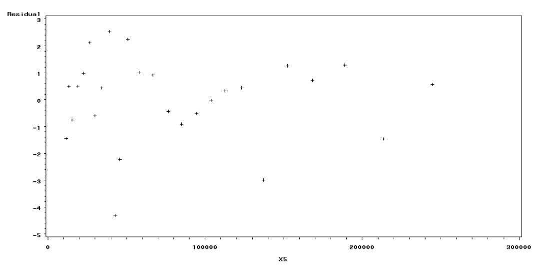
- No variable met the 0.1500 significance level for entry into the model.

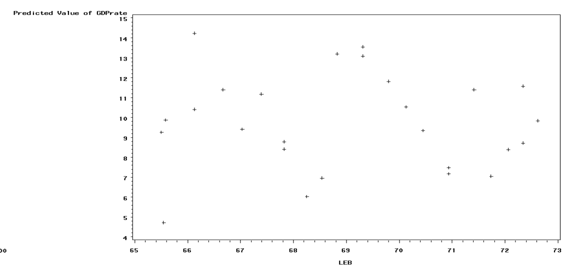


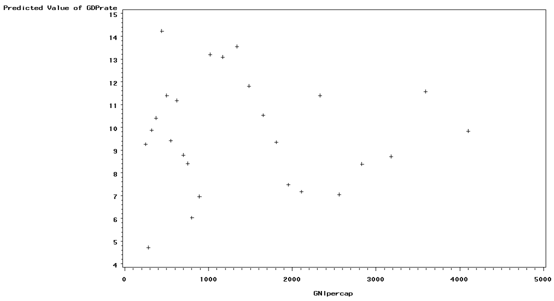
Some plots of residuals versus variables and dependent variable versus independent variables:











By the result of the plots above, the assumption of constant variance does not hold. The variables GNIpercap, X3 and X5 violate the assumption. None of them showed up in my models selected by C(p), so I do not need to worry about them. I found that there are curve relations between predicted GDPrate and LEB, GDPrate and GNIpercap, so I did include those variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. C(p) suggested two models as below:

Number in Model	R-Square	C(p)	Variables in Model
11	0.7496	11.9566	ALP FRT LEB Migration PopGrowth GNIpercap X1 X2 X4 X5 X7
11	0.7483	12.0292	ALP FRT Migration PopGrowth GNIpercap X1 X2 X4 X5 X6 X7

Analysis of Variance

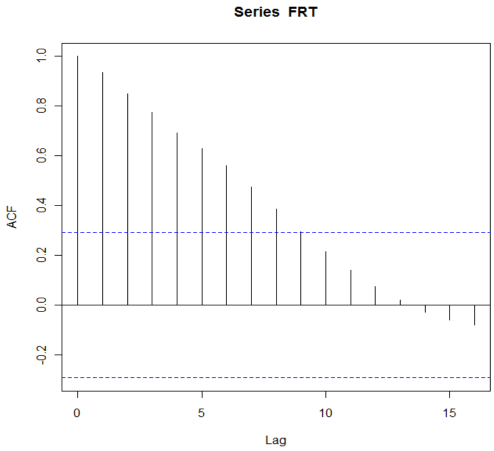
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	134.19776	14.91086	3.21	0.0202
Error	16	74.22570	4.63911		

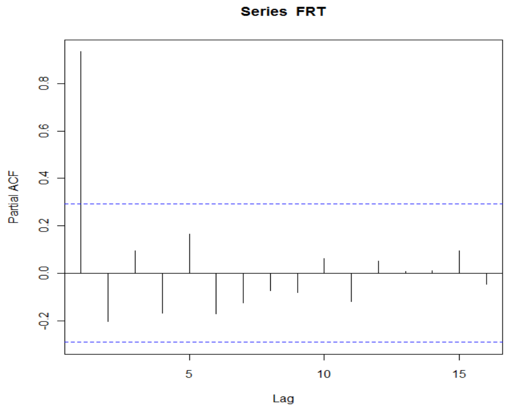
Corrected Total	25	208.42346				
Durbin-Watson D			2.031			
Number of Observations			26			
1st Order Autocorrelation			-0.073			

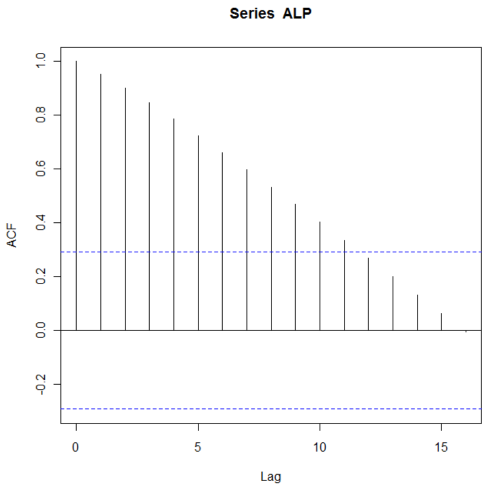
Parameter Estimates

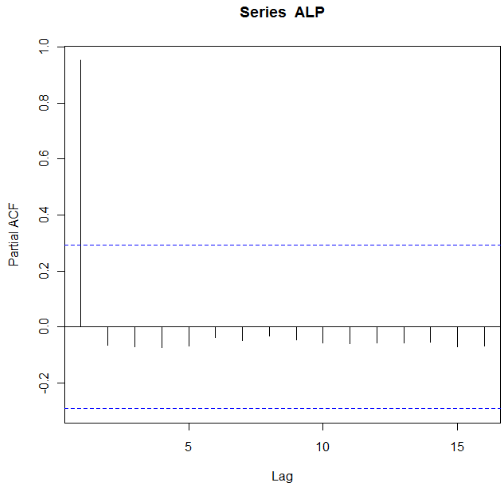
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	921.20660	281.92424	3.27	0.0048
ALP	ALP	1	1.29273	0.55457	2.33	0.0332
FRT	FRT	1	-129.83411	35.35620	-3.67	0.0021
Migration	Migration	1	0.00000408	0.00000606	0.67	0.5103
PopGrowth	PopGrowth	1	-837.07282	350.59100	-2.39	0.0296
X1		1	87.82458	25.45201	3.45	0.0033
X2		1	-0.00000883	0.00000879	-1.00	0.3300
X4		1	9.25498	4.45792	2.08	0.0544
LEB^2		1	-0.14793	0.04492	-3.29	0.0046
GNIpercap^2		1	0.00000134	5.022687E-7	2.67	0.0168

The graphs of time lags for ALP and FRT by ACF and PACF:









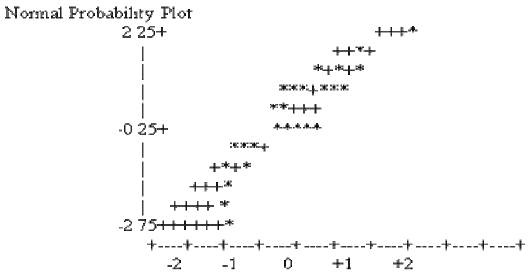
Time lags exist, and lags are 11 for ALP and lags are 8 for FRT.

4.2.12 Egypt

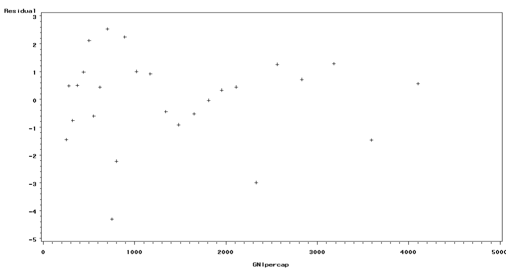
The result of stepwise and the normal plot:

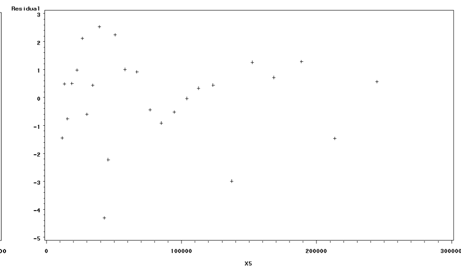
Summary of Stepwise Selection

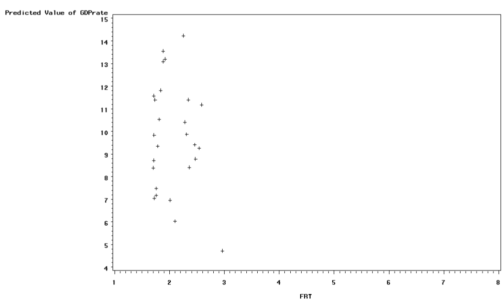
Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	LEB		LEB	1	0.2324	0.2324	8.4787	7.27	0.0126
2	X5			2	0.1325	0.3649	5.2160	4.80	0.0389

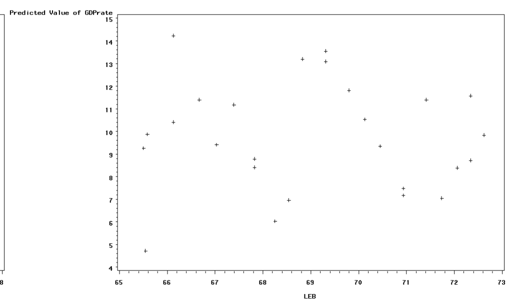


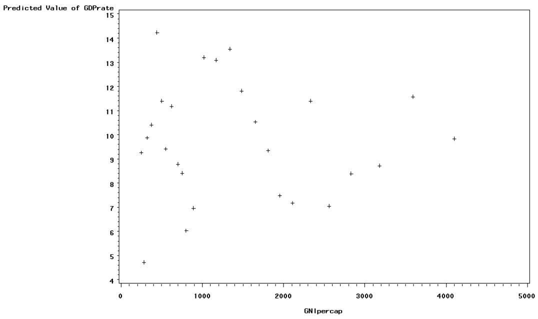
Some plots of residuals versus variables and dependent variable versus independent variables:











By the result of the plots above, the assumption of constant variance does not hold. The variables GNIpercap, X3 and X5 violates the assumption. None of them show up in my models selected by C(p), so I do not need to worry about them. I found that there are curve relations between predicted GDPrate and FRT, GDPrate and LEB, GDPrate and GNIpercap, so I did include those variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. C(p) suggested the model as below:

Number in Model	R-Square	C(p)	Variables in Model
6	**0.5920**	6.9811	LEB Migration X1 X2 X3 X7

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	64.52659	10.75443	4.59	0.0048
Error	19	44.47946	2.34102		
Corrected Total	25	109.00605			

Parameter Estimates

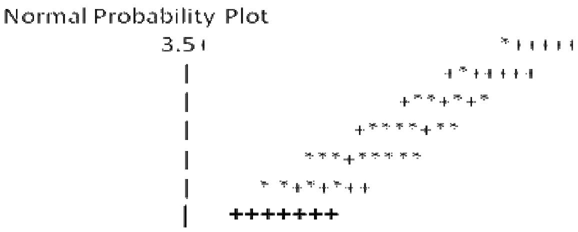
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	**1**	329.54649	96.33637	3.42	0.0029
LEB	LEB	**1**	-10.31876	3.14897	-3.28	0.0040
Migration	Migration	**1**	-0.00006516	0.00002198	-2.96	0.0080
X1		**1**	1.16761	0.60791	1.92	0.0699
X2		**1**	0.00002883	0.00000941	3.06	0.0064
X3		**1**	0.00005111	0.00002042	2.50	0.0216
GNIpercap^2		**1**	0.07786	0.02531	3.08	0.0062
Durbin-Watson D		2.718				
Number of Observations		26				
1st Order Autocorrelation		-0.412				

4.2.13 India

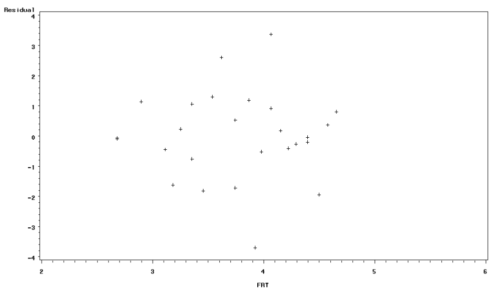
The result of stepwise and the normal plot:

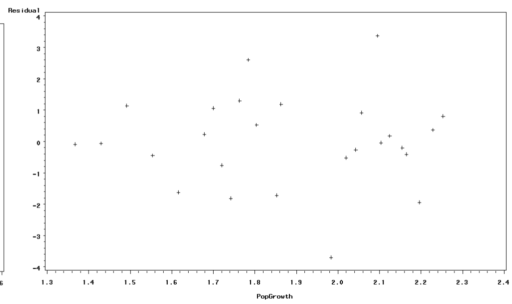
Summary of Stepwise Selection									
Step	**Variable Entered**	**Variable Removed**	**Label**	**Number Vars In**	**Partial R-Square**	**Model R-Square**	**C(p)**	**F Value**	**Pr > F**

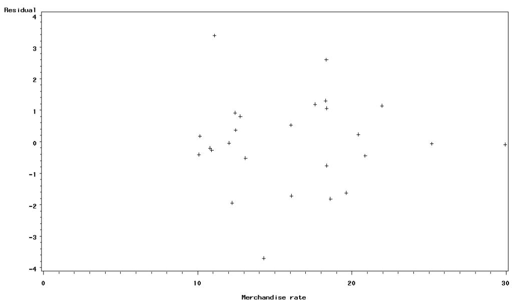
1	ALP		ALP	1	0.1689	0.1689	-1.4469	4.88	0.0370

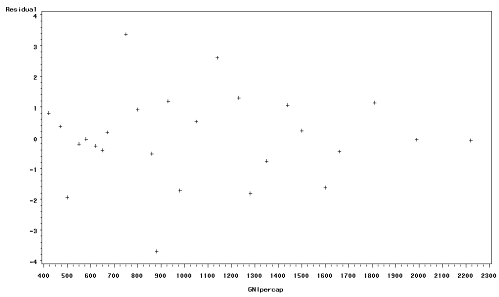


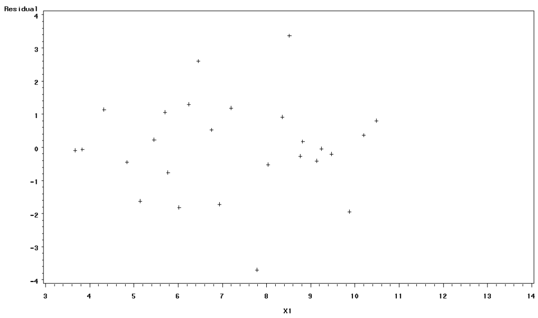
Some plots of residuals versus variables and dependent variable versus independent variables:

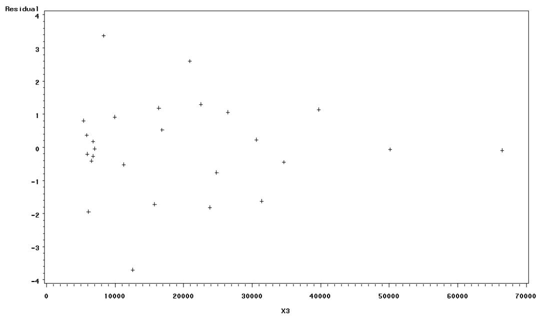


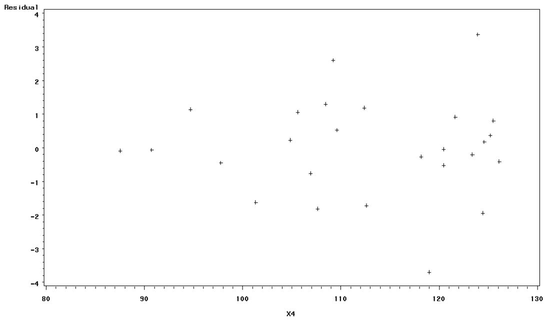


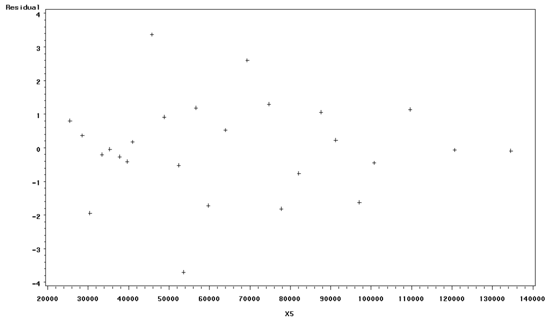


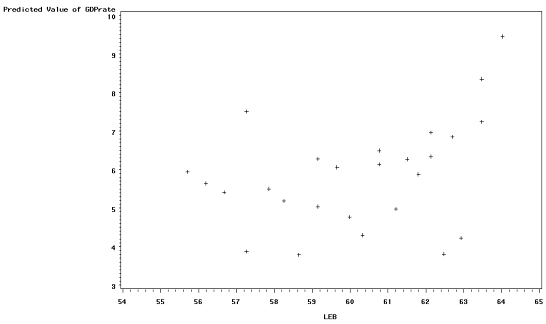


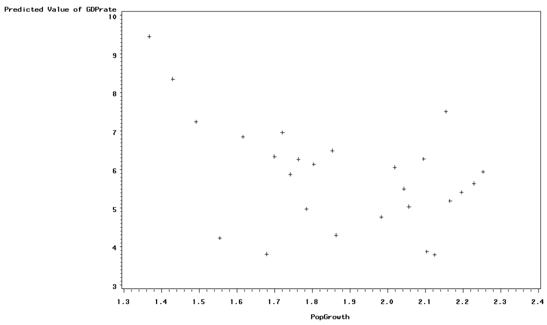












Viewing the result of the plots above, the assumption of constant variance does not holds. The variables FRT, Merchandise_rate, PopGrowth, GNIpercap, X1, X3, X4 and X5 violated the assumption. Then I found that there are curve relations between predicted GDPrate and LEB, predicted GDPrate and PopGrowth, so I did include those variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. But unfortunately, C(p) did give me a model without any significant variables, so I only could run single linear regression model as stepwise suggested.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	28.64325	28.64325	3.26	0.0801

Error	33	289.81796	8.78236		
Corrected Total	34	318.46121			

Durbin-Watson D	2.177
Number of Observations	35
1st Order Autocorrelation	-0.128

Parameter Estimates						
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	Intercept	**1**	-204.86398	116.22821	-1.76	0.0872
ALP	ALP	**1**	3.45654	1.91397	1.81	0.0801

There is no significant variable for India.

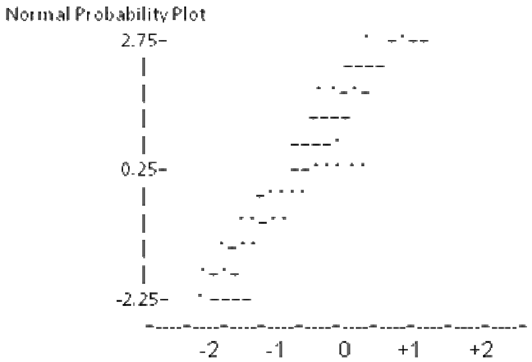
4.2.14 Israel

The result of stepwise and the normal plot:

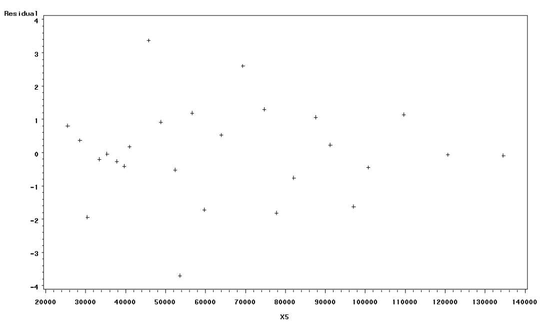
Summary of Stepwise Selection									
Step	**Variable Entered**	**Variable Removed**	**Label**	**Number Vars In**	**Partial R-Square**	**Model R-Square**	**C(p)**	**F Value**	**Pr > F**

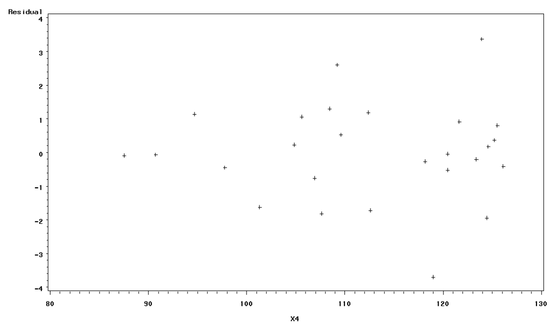
1	X1			1	0.2346	0.2346	14.4208	7.36	0.0122
2	Merchandise_rate		Merchandise rate	2	0.0786	0.3133	12.6788	2.63	0.1182
3	Migration		Migration	3	0.1211	0.4344	8.9160	4.71	0.0411

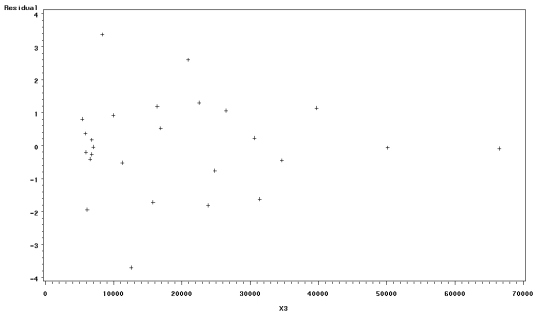
Some plots of residuals versus variables and dependent variable versus independent variables:

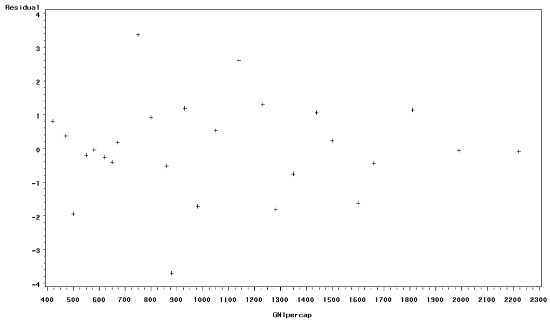


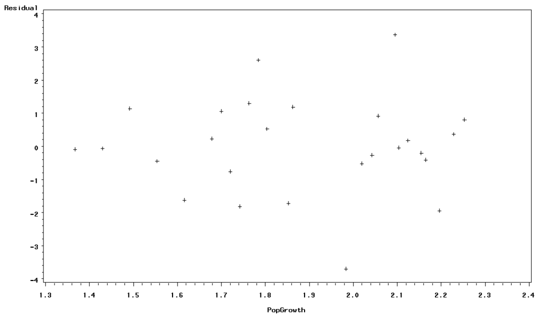
The graphs of time lags for ALP and FRT by ACF and PACF:

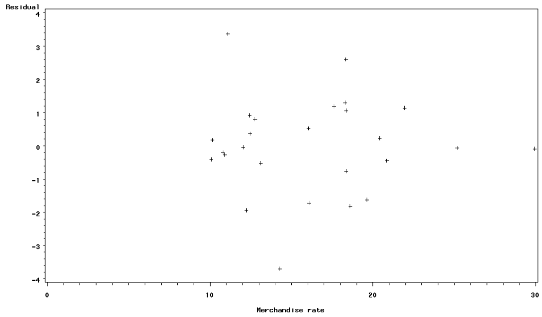


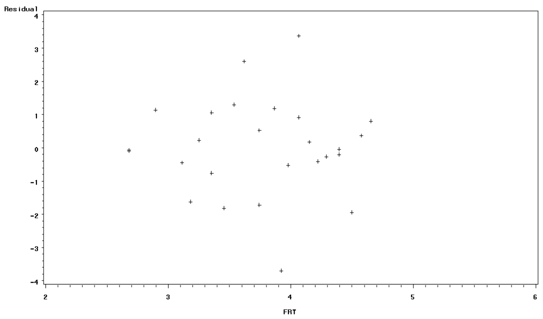












Viewing the result of the plots above, the assumption of constant variance does not holds. The variables FRT, Merchandise_rate, PopGrowth, GNIpercap, X3, X4 and X5 violated the assumption. There is no quadratic form since no plots indicated that. I used C(p) and stepwise to determine a model. C(p) suggested to include the following variables in my regression model:

Number in Model	R-Square	C(p)	Variables in Model
6	**0.6007**	7.0017	ALP GNIpercap X1 X2 X4 X5

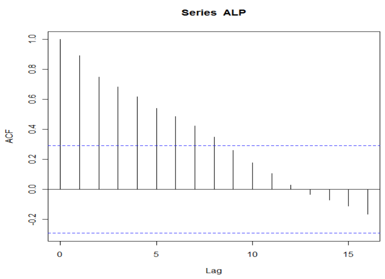
Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	6	99.40332	16.56722	4.76	0.0040
Error	19	66.07853	3.47782	_	_
Corrected Total	25	165.48185	_	_	_

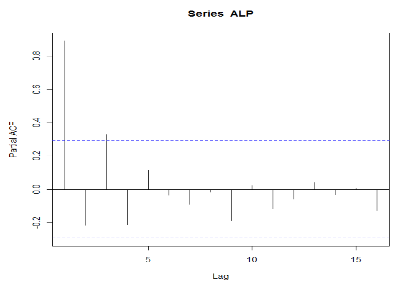
Durbin-Watson D	2.748
Number of Observations	26
1st Order Autocorrelation	-0.380

Parameter Estimates						
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	Intercept	**1**	-216.14230	65.57349	-3.30	0.0038
ALP	ALP	**1**	8.06076	2.42665	3.32	0.0036
GNIpercap	GNIpercap	**1**	0.01263	0.00354	3.57	0.0020
X1	_	**1**	10.33832	3.15410	3.28	0.0040
X2	_	**1**	0.00000326	0.00000102	3.20	0.0048
X4	_	**1**	-0.34518	0.11043	-3.13	0.0056

X5	_	1	-0.00048512	0.00013474	-3.60	0.0019

The graphs of time lags for ALP by ACF and PACF:

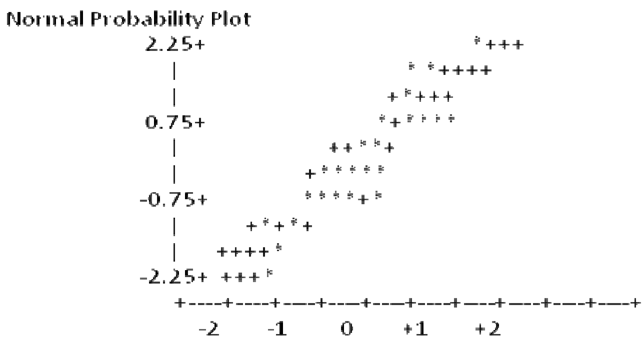




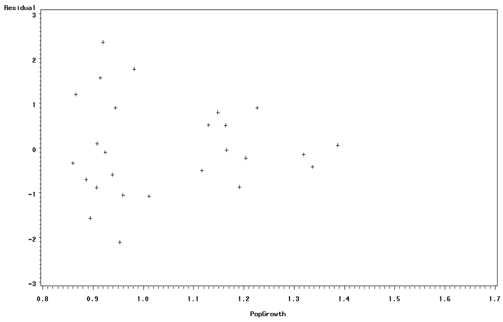
4.2.15 U.S.A.

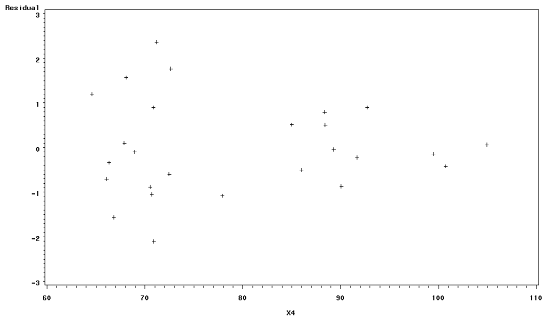
The result of stepwise and the normal plot:

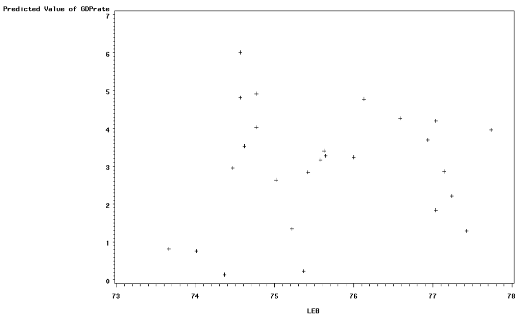
- No variable met the 0.1500 significance level for entry into the model.

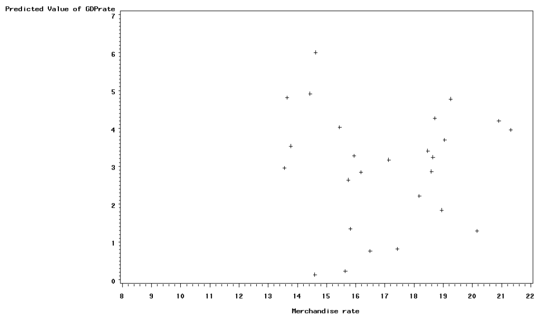


Some plots of residuals versus variables and dependent variable versus independent variables:









Viewing the result of the plots above, the assumption of constant variance does not hold. The variables PopGrowth and X4 violated the assumption. Then I found that there are curve relations between predicted GDPrate and LEB, predicted GDPrate and Merchandise_rate, so I did include those variables' squared term in my regression models. And then I used C(p) and stepwise to determine a model. The models I picked by C(p) are followed by:

Number in Model	R-Square	C(p)	Variables in Model
5	**0.4853**	5.9180	Merchandise_rate Migration GNIpercap X2 X4
5	**0.4846**	5.9467	ALP Migration GNIpercap X2 X4
5	**0.4810**	6.0836	Migration PopGrowth X2 X3 X5

The graphs of time lags for ALP and FRT by ACF and PACF:

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	5	41.85094	8.37019	3.77	0.0144
Error	20	44.38104	2.21905		
Corrected Total	25	86.23198			

Parameter Estimates						
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	Intercept	**1**	56.70898	12.99068	4.37	0.0003
Merchandise_rate	Merchandise rate	**1**	-0.38503	0.25022	-1.54	0.1395
Migration	Migration	**1**	-0.00000817	0.00000214	-3.82	0.0011
GNIpercap	GNIpercap	**1**	0.00014987	0.00007132	2.10	0.0485
X2		**1**	0.00000860	0.00000216	3.98	0.0007
X4		**1**	-0.72041	0.17521	-4.11	0.0005

Durbin-Watson D	2.456
Number of Observations	26
1st Order Autocorrelation	-0.251

Summary of significant variables for the fifteen countries:

Countries	Significant Variables
Argentina	FRT, Merchandise_rate, X3, X5, X7
South Africa	X3
Australia	ALP, Merchandise_rate, GNIpercap, X3, X5, X6
UK	
Brazil	ALP, FRT, GNIpercap, X5
Norway	LEB, GNIpercap, X3
Germany	ALP, Merchandise_rate, GNIpercap, X5
Greece	X3
Italy	
Japan	FRT, X2,
China	ALP, FRT, LEB, PopGrowth, X1, X7
Egypt	LEB, Migration, X2, X3, X7
India	
Israel	ALP, GNIpercap, X1, X2, X4, X5
US	Migartion, GNIpercap, X2, X4

Then I found there is no significant variable for UK, Italy and India; this may be because the data of those three countries is very similar to constant.

4.3 The Mixed Models

Here my general mixed model is:

$$\underline{Y}_i = X_i \underline{\beta} + Z_i \underline{\gamma}_i + \underline{\varepsilon}_i ,$$

$\underline{Y}_i = data\ \ vector\ \ for\ \ the\ \ i^{th}\ \ subject$,

$X_i = known\ \ design\ \ matrix$,

$\underline{\beta} = unknown\ \ fixed\ \ effect\ \ parameters$,

$$Z_i = known\ design\ matrix\ corresponding\ to\ random\ effects\ \underline{\gamma}_i,$$

$$\underline{\gamma}_i = unknown\ random\ effect\ parameters,$$

$$\underline{\varepsilon}_i = vector\ of\ unknown\ residual\ errors,$$

$$\begin{pmatrix} \underline{\gamma}_i \\ \underline{\varepsilon}_i \end{pmatrix} \sim N(\underline{0}, \Sigma_i)$$

$$where\ \Sigma_i = \begin{pmatrix} G & 0 \\ 0 & R_i \end{pmatrix},\ i.e.,\ Var(\underline{\gamma}_i) = G\ \&\ Var(\underline{\varepsilon}_i) = R_i$$

$$and\ \underline{\gamma}_i,\ \underline{\varepsilon}_i\ are\ independent.\ Hence$$

$$\underline{Y}_i \sim N(X_i \underline{\beta}, V_i)\ where\ V_i = Z_i G Z_i' + R_i$$

I fixed all the variables to test if the GDP growth rates in these 15 countries have been significantly different in the last 45 years. In this process, I ran the fixed model with interaction term X3 and X5 first, and then ran without them. I include all the explanatory variables in the fixed model.

4.3.1 Fixed model

For the first model, I fixed all the variables to test if the GDP growth rates in these 15 countries have been significantly different in the last 45 years. Also, I would like to see if any other explanatory variables are significantly effective.

First, I used stepwise to select model including all countries' data and determine if I should include any interaction variable in my fixed model.

Summary of Stepwise Selection

Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F

1	GNIpercap	GNIpercap	1	0.0744	0.0744	19.9491	31.03	<.0001
2	X4		2	0.0108	0.0852	17.2220	4.56	0.0334
3	Merchandise_ra te	Merchandise rate	3	0.0097	0.0949	14.9959	4.11	0.0434
4	FRT	FRT	4	0.0083	0.1032	13.3942	3.52	0.0612
5	LEB	LEB	5	0.0090	0.1121	11.4875	3.85	0.0504
6	ALP	ALP	6	0.0127	0.1248	7.9660	5.51	0.0194
7	PopGrowth	PopGrowth	7	0.0057	0.1304	7.4904	2.48	0.1162

By the result above, I found that X4 as an interaction term is significant when I used stepwise to select model for pooled data set of all countries. So I will run four different fixed models. To distinguish the difference among the four models, I pasted my SAS codes here.

1. With interaction term:

proc mixed data=data; class coutries year;
model GDPrate= year ALP FRT LEB Merchandise_rate Migration PopGrowth GNIpercap X4;
run;

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
year	25	354	1.67	0.0239
ALP	1	354	4.73	0.0304

FRT	1	354	3.83	0.0512
LEB	1	354	4.52	0.0343
Merchandise_rate	1	354	4.21	0.0408
Migration	1	354	0.01	0.9123
PopGrowth	1	354	1.68	0.1953
GNIpercap	1	354	5.02	0.0256
X4	1	354	2.45	0.1182

Fit Statistics	
-2 Res Log Likelihood	1976.5
AIC (smaller is better)	1978.5
AICC (smaller is better)	1978.5
BIC (smaller is better)	1982.4

Covariance Parameter Estimates

Cov Parm	Estimate
Residual	9.4248

2. Without interaction term:

proc mixed data=data; class coutries year;
model GDPrate= year ALP FRT LEB Merchandise_rate Migration PopGrowth GNIpercap; **run;**

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
year	25	355	1.75	0.0155
ALP	1	355	4.86	0.0282
FRT	1	355	7.07	0.0082
LEB	1	355	2.24	0.1351
Merchandise_rate	1	355	5.19	0.0234
Migration	1	355	0.45	0.5006
PopGrowth	1	355	6.29	0.0126
GNIpercap	1	355	12.63	0.0004

Fit Statistics

-2 Res Log Likelihood	1974.7
AIC (smaller is better)	1976.7
AICC (smaller is better)	1976.7
BIC (smaller is better)	1980.6

Covariance Parameter Estimates

Cov Parm	Estimate

Residual	9.4634

Viewing the result above, we do not need to include the interaction term since it is insignificant and also made AIC value bigger than the model without interaction term.

4.3.2 Mixed Model with "Countries" as Random Variable

In the second model, I set "countries" as random variable; the rest of the variables are fixed, because I want to know if the characteristics will remain same for all the countries in the world. Also I repeated year over countries.

1. Compound symmetry

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
year	25	341	2.10	0.0018
ALP	1	341	1.00	0.3172
FRT	1	341	3.75	0.0538
LEB	1	341	8.17	0.0045
Merchandise_rate	1	341	5.77	0.0168
Migration	1	341	0.06	0.8048
PopGrowth	1	341	5.00	0.0259
GNIpercap	1	341	0.01	0.9219

Fit Statistics	
-2 Res Log Likelihood	1870.1
AIC (smaller is better)	1874.1
AICC (smaller is better)	1874.1
BIC (smaller is better)	1875.5

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
CS	coutries	4.4700
Residual		6.4750

2. Unconstructed type

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
year	25	341	2.10	0.0018
ALP	1	341	1.00	0.3172
FRT	1	341	3.75	0.0538
LEB	1	341	8.17	0.0045

Merchandise_rate	1	341	5.77	0.0168
Migration	1	341	0.06	0.8048
PopGrowth	1	341	5.00	0.0259
GNIpercap	1	341	0.01	0.9219

Fit Statistics	
-2 Res Log Likelihood	1870.1
AIC (smaller is better)	1874.1
AICC (smaller is better)	1874.1
BIC (smaller is better)	1875.5

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	coutries	4.4700
Residual		6.4750

3. Heterogeneous compound symmetry

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F

year	25	341	2.70	<.0001
ALP	1	341	1.77	0.1840
FRT	1	341	0.47	0.4922
LEB	1	341	3.86	0.0502
Merchandise_rate	1	341	14.46	0.0002
Migration	1	341	0.04	0.8397
PopGrowth	1	341	2.60	0.1081
GNIpercap	1	341	4.48	0.0351

Fit Statistics

-2 Res Log Likelihood	1819.3
AIC (smaller is better)	1873.3
AICC (smaller is better)	1877.9
BIC (smaller is better)	1892.4

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
Var(1)	coutries	5.2544
Var(2)	coutries	10.8078
Var(3)	coutries	12.0020
Var(4)	coutries	17.0472

Var(5)	coutries	14.1423
Var(6)	coutries	22.7866
Var(7)	coutries	12.1580
Var(8)	coutries	11.3359
Var(9)	coutries	18.4742
Var(10)	coutries	18.9645
Var(11)	coutries	14.5672
Var(12)	coutries	24.7726
Var(13)	coutries	20.7147
Var(14)	coutries	10.8439
Var(15)	coutries	5.9191
Var(16)	coutries	7.2517
Var(17)	coutries	4.1089
Var(18)	coutries	6.9435
Var(19)	coutries	7.6519
Var(20)	coutries	9.5835
Var(21)	coutries	6.3913
Var(22)	coutries	9.5947
Var(23)	coutries	24.0114
Var(24)	coutries	10.7198
Var(25)	coutries	6.3711
Var(26)	coutries	7.0801
CSH	coutries	0.4960

4. Randomized countries

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
year	25	341	2.10	0.0018
ALP	1	341	1.00	0.3172
FRT	1	341	3.75	0.0538
LEB	1	341	8.17	0.0045
Merchandise_rate	1	341	5.77	0.0168
Migration	1	341	0.06	0.8048
PopGrowth	1	341	5.00	0.0259
GNIpercap	1	341	0.01	0.9219

Fit Statistics

-2 Res Log Likelihood	1870.1
AIC (smaller is better)	1874.1
AICC (smaller is better)	1874.1

BIC (smaller is better)		1875.5
Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
coutries		4.4700
year	coutries	6.4750

Now let's compare the four models above to determine an appropriate one as my mixed model to predict if culture has significant effect on economic growth.

Model Name	**AIC**	**BIC**	**Significant Variables**
Compound Symmetry	1870.1	1875.5	Year, LEB, PopGrowth, Merchandise_rate
UN	1870.1	1875.5	Year, LEB, PopGrowth, Merchandise_rate
Heterogeneous Compound Symmetry	1819.3	1892.4	Year, LEB, PopGrowth, Merchandise_rate
random countries	1870.1	1875.5	Year, LEB, PopGrowth, Merchandise_rate

By comparing the AIC and BIC, I found that the third model is the best one. All the models have the same significant variables year, LEB, PopGrowth and Merchandise_rate. So I determine the third one as my mixed model.

Chapter 5: Conclusion

First, there is no evidence to show that the countries in the same cluster with similar cultural features will also have the same significant variables on their economic growth. This is proved by connecting the clusters results to multiple regression models. For example, Argentina and South Africa are in the same cluster, but the variables are significant in regression model for Argentina but not for South Africa. The smallest distance cluster is the one between Italy and Japan; although this means they have very similar culture features, they still have completely different significant variables in their own regression models respectively. There is no significant variable for Italy. The other clusters have the same situation. So although the countries have similar culture features, the ways culture raises GDP growth are different. It may be because the countries with similar culture features have different economic institutions and political policies which produce some restriction on effect of culture.

Second, by the results of multiple linear regressions of different countries we have enough evidence to say that culture has causal relation to economic growth. We can just conclude that culture has significant effect on economic growth and development for those fifteen countries. But there are still three countries: U.K, Italy and India, which have no significant variable, and also, several countries only have one significant variable. This lessens the statement that culture has significant effect on economic growth. So why can this happen? Because we lack of some important data such as education level, ration of boys and girls, and proportion of population in urban of total population in a country. Also, I have only one dependent variable—GDP growth rate--as a measurement tool of economic growth. But the results of economic growth and development also include some other factors such as the satisfaction of national populace, multiple strength of a country, people's living pressure or happiness indices or people's average income. So because of the uncertainties, we are still in decision of the conclusion of this part.

Finally, the outputs of mixed models tell us some factors have significant effect on GDP growth. But what I am looking for is that the factors do include culture or some parts of culture. Then I found in both models (one is fixed and the other is mixed), Merchandise_rate and PopGrowth are always significant variables, which means over the different levels of PopGrowth and Merchandise_rate, GDP rate has significant difference over the 45 years. That means population growth rate and merchandise annual rate have significant effect through some factors that cause increase in GDP growth. I think PopGrowth may affect economic growth by changing the population that is the source of labor and human resources. Merchandise_rate as a significant variable can be explained by that this variable's effects on economic growth by international trade. The variables ALP, FRT and GNIpercap are significant in the fixed model and the variables LEB is significant in the mixed model. Within the 15 countries ALP is significantly effective but insignificant worldwide, I think it may be because agricultural products do not contribute much to economic growth worldwide, but they does in these 15 countries I selected. The same thing happened to GNIpercap. It may be because GNI per capital is very small in most countries of the world, but it is high in these 15 countries. FRT relates to population growth, so it may affect economic growth by changing the population that is the source of labor and human resources. LEB as a significant variable can prove that the expectation of how many years a person can live at the birth affects economic growth. Because LEB relates to people's attitude towards future life, the expectance of future economic growth development and stability of political structure and the improvement of growth environment, also it reviews historical people's longevities and economic growth conditions and the current economic conditions, can affect people's behavior in economic activities. Thus it also can imply that culture has effect on economic growth.

In summary, we found although there is no obvious evidence that can directly and straightly prove culture causes economic growth, we can still conclude that culture can affect economic growth and development.

References

James Bullard, Jasmina Arifovic, John Duffy (1995): Learning in a Model of Economic Growth and development. *Journal of Economic Growth. June 1997*

Guido Tabellini (2006): Culture and institutions: economic development in the regions of Europe. *ftp://ftp.igier.uni-bocconi.it/wp/2005/292.pdf*

Guido Tabellini (2007): Institutions and Culture. *http://www.cesifo-group.de/DocCIDL/cesifo1_wp1492.pdf*

Avner Greif and Guido Tabellini (2009): Cultural and Institutional Bifurcation: China and Europe Compared. *http://hdl.handle.net/10.1257/aer.100.2.135*

Daron Acemoglu, Simon Johnson, James Robinson (2004): Institutions as the Fundamental Cause of Long-Run Growth. *http://elsa.berkeley.edu/~chad/handbook9sj.pdf*

F. Fukuyama: Culture and Economic Development: Cultural Concerns. *June 2010, Vol. 8, No. 4, Pages 677-716*

John Luke, Gallup Jeffrey D. Sachs, Andrew D. Mellinger (1999): Geography and Economic Development. *http://www.cid.harvard.edu/cidwp/pdf/001.pdf*

Keith Griffin (1997): Culture, Human Development and Economic Growth. *http://www.unrisd.org/unrisd/website/document.nsf/(httpPublications)/A96A30394656266E80256B67005B6D37?OpenDocument*

Herb Thompson (2001): Culture and Economic Development: Modernisation To Globalisation. *http://theoryandscience.icaap.org/content/vol002.002/thompson.html*