Inequality amongst the wealthiest
and its link with economic growth

Philip Hans Franses

Stephanie Vermeer

Econometric Institute

Erasmus School of Economics

Econometric Institute Report 2010-27

Abstract

In this paper we correlate the key features of the distribution of wealth of the 500 wealthiest individuals in the Netherlands with economic growth and stock market returns for 1998 to 2009. We show that each year the distribution obeys a power law and that the key parameter measures the degree of inequality. Our main finding is that more inequality amongst the wealthiest is associated with higher economic growth.

JEL code: D31, E21

Key words: Wealth distribution; Power law; Economic Growth

This version: April 23 2010

1

1. Introduction

Each year in the fall, the Netherlands-based magazine Quote publishes its Quote 500 ranking of the 500 wealthiest families and individuals. They started this publication in 1998, so up till now there are 12 such rankings including 2009. In this paper we study the distributional properties of these rankings, and we examine the evolution of these over time. To the best of our knowledge, these Quote 500 data have never been analyzed before. Additionally, and this is new to the literature, we correlate the properties of these distributions with real economic growth and with stock market data.

This paper is largely of an exploratory nature. We describe the features of the data, and we do so using standard and a few new techniques. One new method involves the clustering of the 500 individual ranks into a smaller set of ranks that are associated with approximately similar wealth levels. For example, we will show that for some years the 500 individual quotes can be summarized by 6 clusters with similar quotes. When doing so, we can for example conclude that the cluster with individuals with the largest wealth gets smaller over time.

The outline of our paper is as follows. In Section 2 we provide basic statistics of the wealth rankings over the years. Interestingly, the data obey so-called power laws almost perfectly. We also examine the evolution of the basic statistics over time. In Section 3 we correlate a few of the main features with real economic variables as Gross Domestic Product (GDP) growth and stock market returns. We find that more economic growth leads to more inequality amongst the wealthiest. In Section 4 we conclude this paper with a discussion of how our findings fit with the findings presented in the relevant literature.

2. Descriptive statistics of the wealth data

Each year since 1998 the magazine Quote publishes its ranking of the 500 wealthiest families or individuals in the Netherlands. Ever since 1998 the Royal Family is present in the top 10, and also well-known families like Heineken and Fentener van Vlissingen are

included. The way the list is compiled has been rather constant over time, so it is possible to evaluate any tendencies over time.

Basic statistics

We created a database with all the 500 entries each year, and in total there are more than 900 entries over the 12 years of data, covering 1998 to and including 2009. Hence, each year there are new entries and there are exists. Sometimes individuals disappear from the list for a while, and then re-enter a few years later. We recalculated the wealth figures in 1998 to 2001 in terms of euros, as before 2002 the figures were given in guilders.

Insert Figure 1 about here
Insert Table 1 about here

In Table 1 we give a few basic summary statistics of the distribution for each of the twelve years. As is also obvious from the graph in Figure 1, for the data for 2004, the distributions of the wealth data are extremely skewed. The means and the medians in Table 1 are very different, and also the difference between the maximum values and the minimum values is large. In the period 1998-2009 there are two episodes with less than average economic growth, and these are around 2001-2003 and 2009. These episodes are also reflected by the numbers in Table 1, where the mean, median, minimum and maximum are clearly smaller in those periods than in other periods. The maximum value across all years and individuals is observed for 2008 with a value of 24100 million euros. To give an impression of the sheer size of this wealth of about 24 billion euros, it is good to know that annual GDP in the Netherlands in the observed time frame is around 400 to 500 billion euros. Finally, Table 1 shows that the distribution of wealth is not constant over the years, and major changes in means, medians and standard deviations can occur.

Insert Table 2 about here

The changes in the distributions can also be observed from the numbers in Table 2, which concern the individual changes. For example, in 1999 there are 418 families and individuals who were also on the list in 1998, and the mean percentage change in their wealth is 27.8%. The median is usually smaller than the mean, which here says that there is a distribution skewed to the right. Indeed, each new year more individuals have more wealth and those who substantially lost their wealth will more likely drop out from the list. Table 2 further shows that the individual differences can be quite large. Also, entry and exit is in between 25 to 50, with the largest dropout in 1999, 2000 and 2001, which may have to do with the dotcom crisis in those years. Finally, the exceptional crisis year of 2009 is reflected by an average decrease in wealth of about 9%.

Insert Table 3 about here
Insert Figure 2 about here

To get a first impression of how the distributions may change over time, we run a regression of the individual wealth in year T on that in year T-1. The number of data points is the same as those reported in Table 2, and hence the regressions include 418 to 476 observations for 1998 to 2009, respectively. The results are presented in Table 3, and it is clear that the distributions change over time. Indeed, when the distributions would be most similar, then the intercept should be zero and the slope parameter should be equal to one. Comparing the estimates in Table 3 with their standard errors, we see that only for the years 2006 and 2007 the distributions do not differ much from those in the years before. The final column gives the R^2 for each of the regressions, and clearly the fit is quite high. Hence, there seems to be some information in year T-1 that can be used to forecast the data in year T. A graphical illustration of this feature in presented in Figure 2, where the regression line approximates the 45 degree line.

Power laws

An alternative way to summarize the data amounts to fitting a distribution to the data. In other studies where wealth data are analyzed empirically, it has been documented that

such data can well be described by a Pareto distribution or more in general, by a power law, see for example Persky (1992) and Davies and Shorrocks (2000). Castaldi and Milakovic (2007) study the Forbes 400 list (which lists the 400 wealthiest American families and individuals) and they find supportive evidence for the presence of a power law, see also Levy and Solomon (1997). In brief, a power law implies that there is a link between the wealth value and the rank on the list. The (slightly more strict) law of Zipf states that the first ranked has twice as much wealth as the second ranked, and that this second ranked has twice as much wealth as the number four on the list. Mathematically this implies that there is a linear link between the (natural) logarithm of wealth and the (natural) logarithm of the rank. A very readable survey of power laws is given in Newman (2005).

Insert Figure 3 about here
Insert Table 4 about here

In Figure 3, which consists of six panels, each covering two years of data, we depict the correlation between the log of wealth and the log of the rank. The graphs are quite striking in the sense that the scatters indeed suggest linear links for all years. We also see that the distributions change over time, as sometimes the lines do not overlap.

To quantify the precise relation, we regress the log of wealth on the log of the rank for each of the twelve years, and the main results are presented in Table 4. Of course, the regression line would be the least precise in the upper tail of the distribution, see the discussion on the proper estimators in Castaldi and Milakovic (2007), but for our purposes the least squares estimate of the regression line will suffice. The estimates of the slope are all around -1, with the largest absolute values in 1998, 1999 and 2006, and the smallest absolute values in 2001-2004. The R^2 values of the regression line are all very close to 1, and hence the fit is very good. So, we see that also the Quote 500 distributions closely follow a power law. Note that the larger the absolute value of the slope is, the larger is the inequality of wealth in the sense that the difference between top values and bottom values of wealth is larger. As could be expected from the various panels of Figure 3, the power law holds for all years, even though the distributions changes over time.

Clusters of ranks

The graphs in Figure 2 show that the distribution of points is not uniform over the line, that is, there are more wealth data in between, say, 50 to 100 million euros than that there are in between 200 and 500 million euros. It may thus be that the 500 entries of the distribution can be summarized into a smaller set of ranks where for each new rank there is a cluster of families and individuals with approximately the same wealth.

Insert Figure 4 about here

To examine the presence of clusters, we take a closer look at the distribution of the data, after taking natural logarithms. In Figure 4 we present a typical graph of such a distribution. Clearly, the empirical distribution shows multi-modality. We exploit this feature by fitting a mixture of S normal distributions to the data. In brief, for each of the 500 data points we assign a probability p_1 that it is a draw from a normal distribution with mean μ_1 and variance σ^2, a probability p_2 that it is a draw from such a distribution with mean μ_2 and variance σ^2, and so on. We restrict the variance of these distributions to be equal in order to be able to fit a distribution to clusters with just a few observations. Based on the estimated probabilities for each observation, we use the 0.5 cut-off point to assign an observation to one of the S clusters. Some experimentation indicated that a maximum value of 6 for S should do. We use the Akaike and Bayesian information criteria (AIC and BIC) to make a choice for S.

Insert Tables 5 and 6 about here

In Table 5 we give the AIC values and in Table 6 the BIC values, where the smallest values are in italics. Interestingly, for both criteria we see that the preferred value of S is either 5 or 6, and never below 5. When AIC and BIC are in agreement, we take that particular value of S, and when the criteria disagree we follow BIC.

Insert Table 7 about here

In Table 7 we give the estimation results for each of the twelve years. A first observation is that the estimate of the variance σ^2 is remarkably constant over the years. Second, the standard errors for the μ_s parameters are quite small, so apparently the S clusters are rather distinctive. Only in case a cluster contains a single observation (2005, 2008, and 2009) the standard error is large due to lack of data. Note that the estimates for the means concern log-transformed data. When the cluster with the highest mean value gets the new rank 1 and the next cluster rank 2 and so on, one could graphically show that again there is a linear link between these means and the log of the rank. Hence, the power law is robust to clustering.

Insert Table 8 about here

In Table 8 we give the ranks out of the 500 that are now clustered in the new ranks 1 to 5 (or 6). Additionally, we give the range of the wealth in each of these clusters. We learn from Table 8 that the cluster with families and individuals with the largest wealth gets smaller in size over time. In 1998 and 1999 it contained 9 and 8 observations, while in the last two years it consisted of just 1 observation. Furthermore, the cluster with the smallest size of wealth is on average of size 300, meaning that about 300 of the 500 entries in the Quote 500 ranking approximately have similar wealth and are difficult to distinguish from each other. The second wealthiest cluster is of size 10 to 20, and this size is rather constant over time. In sum, the Quote 500 list can be sensibly summarized in, what could be called, a Quote 6 list, which consists of six clusters with approximately similarly wealthy families and individuals. An implication is that when individuals have higher rankings over time, it really matters (in a statistical sense) if they enter a higher cluster.

In sum, the Quote 500 rankings of wealthy families and individuals change over time, although each year's distribution can neatly be approximated by a power law. When the rankings are summarized into a ranking of six clusters, the power law is preserved. It would now be interesting to see of the distribution of wealth is somehow correlated with the real economy.

3. Correlating wealth with the economy

Now we have characterized the distribution of the Quote 500 data over the years, it is of interest to examine if the key features if these distributions have something in common with the real economy. As key variables that could be associated with wealth we use real GDP growth and the (annual) stock market returns on the Amsterdam Stock Exchange (AEX). The data for the Netherlands are given in Table 9.

Insert Table 9 about here

Insert Figure 5 about here

Some unreported analysis of the data indicates that stock market returns do not correlate with wealth, that is, after correction for effects of economic growth these returns do not add much information. In fact we find that one-year lagged returns have predictive value for economic growth. So, the rest of the analysis now purely focuses on economic growth.

First, when we regress economic growth on the mean and median of the changes in wealth (Table 2), and estimate the parameters using Two-Stage Least Squares (TSLS) to account for simultaneity (with one-year-lagged variables as instruments, we get slopes of 0.190 (0.072) and 0.394 (0.066), respectively, with standard errors in parentheses. In words, more economic growth is associated with higher increases in wealth, see also Figure 5.

Second, when we analyze the links between the α coefficients in Table 4 and economic growth, we find the growth negatively correlates (-25.6 with 12.11 as standard error) with the current value of the coefficient and that current α negatively correlates (-0.020 with standard error 0.008) with lagged growth. In words, more inequality in wealth is associated with higher economic growth, while past positive economic growth makes inequality to increase further.

This last finding is further supported when we regress the logs of the means, medians and standard deviations in Table 1 on lagged economic growth, where each of the estimated slopes is significant and positive. If we were to use these last three models

for forecasting the properties of the Quote 500 distribution in 2010, we would predict a mean of 89.8, and median of 39.0 and a standard deviation of 264.9.

4. Conclusions

In this paper we have analyzed the distributional properties of the Quote 500 data on the 500 wealthiest families and individuals in the Netherlands. It turns out that the distribution changes over time, and that these changes are correlated with economic growth. Even though there are changes in the distribution, we find that the wealth data can be adequately described by a power law. Such a power law is also found for the data when they are clustered in six clusters with each approximately similarly wealthy families and individuals.

A main finding in our paper is that lagged economic growth causes inequality amongst the wealthiest to increase. In the literature on economic growth and inequality this is not a common finding. In fact, usually the correlation is found the other way around, that is, economic growth decreases inequality and also less inequality makes economies to grow. A contrasting argument is presented in Barro (2000), where he argues that the positive correlation could well exist if one were to look only at richer countries. Hence, our finding supports his line of thought that when there is already a base level of income or wealth, that then a further increase in economic output would even increase inequality.

Table 1:

Summary statistics of the wealth of the 500 wealthiest in the Netherlands (in millions of euros)

Year	Mean	Median	Minimum	Maximum	Standard Deviation
1998	256.381	82.5	35	8500	764.576
1999	310.830	110	45	15000	1005.332
2000	362.500	140	60	15000	1075.399
2001	362.960	150	75	17000	1011.177
2002	165.316	68	32	8000	462.294
2003	180.524	73	35	10000	530.499
2004	194.850	80	41	12000	602.847
2005	219.860	86	43	12500	642.319
2006	241.550	91	46	15000	767.987
2007	273.558	100.5	48	18500	934.301
2008	289.030	110.0	50	24100	1133.854
2009	254.012	100	45	20500	968.731

Table 2:

Summary statistics of the annual percentage change in individual wealth

Year	Individuals	Mean	Median	Minimum	Maximum
1999	418	27.768	11	-75	614
2000	431	25.188	8	-69	447
2001	426	11.322	6	-77	450
2002	452	1.431	-2	-70	512
2003	454	9.753	6	-64	251
2004	449	7.310	4	-67	172
2005	459	11.020	4	-57	176
2006	466	7.129	5	-65	143
2007	475	10.263	6	-66	164
2008	473	10.460	5	-62	658
2009	476	-9.189	-12	-89	160

Table 3:

A regression of log(wealth in year T) on intercept and log(wealth in year T-1), with estimated standard errors in parentheses

| | | Parameters | | |
Year	Number of cases	Intercept	Log(wealth in year T-1)	R^2
1999	418	0.552 (0.085)	0.921 (0.017)	0.873
2000	431	0.505 (0.085)	0.932 (0.017)	0.881
2001	426	0.621 (0.086)	0.892 (0.016)	0.877
2002	452	-0.650 (0.079)	0.968 (0.015)	0.906
2003	454	0.202 (0.055)	0.970 (0.012)	0.936
2004	449	0.145 (0.042)	0.981 (0.009)	0.965
2005	459	0.137 (0.049)	0.989 (0.010)	0.953
2006	466	0.048 (0.043)	1.001 (0.009)	0.965
2007	475	0.074 (0.046)	1.002 (0.009)	0.960
2008	473	0.278 (0.061)	0.955 (0.012)	0.930
2009	476	0.096 (0.063)	0.957 (0.012)	0.926

Table 4:

Parameter estimates (and standard errors) for power laws. Based on the regression

$$\log(\textit{Wealth}) = \mu + \alpha \log(\textit{Rank}) + \varepsilon$$

Year	$\hat{\mu}$	$\hat{\alpha}$	R^2
1998	10.092 (0.022)	-1.038 (0.004)	0.991
1999	10.080 (0.020)	-0.988 (0.004)	0.993
2000	10.112 (0.016)	-0.952 (0.003)	0.995
2001	9.916 (0.013)	-0.894 (0.002)	0.996
2002	9.150 (0.013)	-0.898 (0.002)	0.997
2003	9.224 (0.015)	-0.894 (0.003)	0.995
2004	9.237 (0.018)	-0.879 (0.003)	0.993
2005	9.453 (0.018)	-0.899 (0.003)	0.993
2006	9.559 (0.016)	-0.980 (0.003)	0.995
2007	9.743 (0.017)	-0.926 (0.003)	0.994
2008	9.737 (0.027)	-0.908 (0.007)	0.984
2009	9.575 (0.027)	-0.899 (0.005)	0.984

Table 5:

AIC for mixtures of normal distributions for log(wealth) data

Mixtures	2	3	4	5	6
Year					
1998	2.634	2.557	2.508	2.502	*2.455*
1999	2.563	2.483	2.404	2.384	*2.368*
2000	2.453	2.372	2.339	2.280	*2.270*
2001	2.317	2.222	2.184	*2.151*	2.153
2002	2.311	2.186	2.141	2.110	*2.105*
2003	2.306	2.159	2.103	2.093	*2.073*
2004	2.280	2.094	2.036	2.102	*1.992*
2005	2.319	2.178	2.116	2.071	*2.065*
2006	2.346	2.191	2.144	2.080	*2.076*
2007	2.406	2.277	2.211	2.153	*2.140*
2008	2.438	2.335	2.275	2.192	*2.185*
2009	2.419	2.329	2.256	2.227	*2.206*

Table 6:

BIC for mixtures of normal distributions for log(wealth) data

Mixtures	2	3	4	5	6
Year					
1998	2.668	2.608	2.576	2.587	*2.556*
1999	2.597	2.534	2.471	*2.468*	2.469
2000	2.487	2.422	2.406	*2.365*	2.371
2001	2.351	2.272	2.252	*2.235*	2.254
2002	2.344	2.237	2.208	*2.194*	2.206
2003	2.339	2.210	2.179	2.177	*2.174*
2004	2.314	2.145	2.103	2.096	*2.092*
2005	2.352	2.228	2.193	*2.155*	2.166
2006	2.380	2.242	2.212	*2.164*	2.177
2007	2.440	2.328	2.278	*2.237*	2.241
2008	2.472	2.385	2.343	*2.277*	2.286
2009	2.453	2.380	2.323	2.311	*2.307*

Table 7:
Parameter estimates (and standard errors) for the clusters

Year	$\hat{\sigma}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_5$	$\hat{\mu}_6$
1998	0.270	3.854	4.597	5.418	6.420	7.375	8.530
	(0.016)	(0.031)	(0.040)	(0.050)	(0.061)	(0.166)	(0.079)
1999	0.333	4.277	5.254	6.318	7.238	8.762	
	(0.020)	(0.029)	(0.046)	(0.095)	(0.160)	(0.090)	
2000	0.312	4.518	5.395	6.356	7.669	9.212	
	(0.018)	(0.030)	(0.047)	(0.067)	(0.062)	(0.161)	
2001	0.309	4.710	5.511	6.397	7.624	9.196	
	(0.021)	(0.029)	(0.063)	(0.073)	(0.060)	(0.133)	
2002	0.300	3.935	4.728	5.584	6.815	8.410	
	(0.018)	(0.026)	(0.067)	(0.064)	(0.062)	(0.118)	
2003	0.256	3.971	4.559	5.356	6.158	7.039	8.725
	(0.018)	(0.028)	(0.061)	(0.052)	(0.092)	(0.059)	(0.096)
2004	0.235	4.059	4.602	5.391	6.088	7.080	8.767
	(0.016)	(0.027)	(0.051)	(0.044)	(0.070)	(0.054)	(0.062)
2005	0.304	4.255	5.301	6.229	7.289	9.431	
	(0.016)	(0.020)	(0.046)	(0.092)	(0.090)	(>100)	
2006	0.300	4.308	5.349	6.203	7.306	9.062	
	(0.016)	(0.020)	(0.051)	(0.089)	(0.072)	(0.115)	
2007	0.306	4.367	5.405	6.373	7.541	9.256	
	(0.015)	(0.021)	(0.048)	(0.065)	(0.092)	(0.118)	
2008	0.287	4.406	5.410	6.324	7.413	10.091	
	(0.013)	(0.021)	(0.039)	(0.060)	(0.058)	(>100)	
2009	0.266	4.263	5.074	5.816	6.687	7.493	9.927
	(0.016)	(0.021)	(0.052)	(0.057)	(0.094)	(0.075)	(>100)

Table 8:
Ranks of the data based on clusters, part I

Rank based on clusters	Rank in 1-500	Wealth (x Million Euros)
	1998 (S = 6)	
1	1-9	8500-3000
2	10-17	1900-1200
3	18-55	1000-400
4	56-128	360-160
5	129-298	150-70
6	299-500	67.5-35
	1999 (S = 5)	
1	1-8	15000-3400
2	9-20	2200-1100
3	21-69	1000-370
4	70-220	360-130
5	221-500	125-45
	2000 (S = 5)	
1	1-4	15000-6500
2	5-25	4000-1200
3	26-75	1000-400
4	76-222	385-155
5	223-500	150-60
	2001 (S = 5)	
1	1-3	17000-6000
2	4-26	4100-1200
3	27-76	1000-425
4	77-200	400-185
5	201-500	180-75

Rank based on clusters	Rank in 1-500	Wealth (x Million Euros)
	2002 (S = 5)	
1	1-3	8000-2600
2	4-25	1700-600
3	26-81	500-190
4	82-185	185-86
5	186-500	85-32
	2003 (S = 6)	
1	1-2	10000-3800
2	3-21	2300-780
3	22-43	726-355
4	44-118	348-148
5	119-228	146-79
6	229-500	78-36
	2004 (S = 6)	
1	1-2	12000-3400
2	3-22	2000-800
3	23-47	675-350
4	48-128	332-152
5	129-237	150-83
6	238-500	82-41
	2005 (S= 5)	
1	1	12500
2	2-21	3500-945
3	22-55	850-365
4	56-163	350-135
5	164-500	130-43

Rank based on clusters	Rank in 1-500	Wealth (x Million Euros)
	2006 (S = 5)	
1	1-2	15000-4900
2	3-22	2700-972
3	23-58	866-367
4	59-167	359-139
5	168-500	136-46
	2007 (S = 5)	
1	1-2	18500-5600
2	3-16	3900-1200
3	17-62	1100-400
4	63-180	390-145
5	181-500	142-48
	2008 (S = 5)	
1	1	24100
2	2-23	4100-1000
3	24-74	950-384
4	75-201	377-150
5	202-500	147-50
	2009 (S = 6)	
1	1	20500
2	2-13	3600-1300
3	14-37	1200-600
4	38-109	550-246
5	110-223	235-115
6	224-500	113-45

Table 9:

Annual economic growth and stock market returns for the Netherlands

	Real GDP growth	Stock market return
1997	4.3	40.79
1998	3.9	29.99
1999	4.7	24.71
2000	3.9	-5.04
2001	1.9	-20.52
2002	0.1	-36.32
2003	0.3	4.62
2004	2.2	3.09
2005	2.0	25.48
2006	3.4	13.41
2007	3.6	4.12
2008	2.0	-52.32
2009	-4.0	36.35

Series: Y2004	
Sample 1 500	
Observations 500	
Mean	194.8500
Median	80.00000
Maximum	12000.00
Minimum	41.00000
Std. Dev.	602.8466
Skewness	15.63093
Kurtosis	297.4783
Jarque-Bera	1826974.
Probability	0.000000
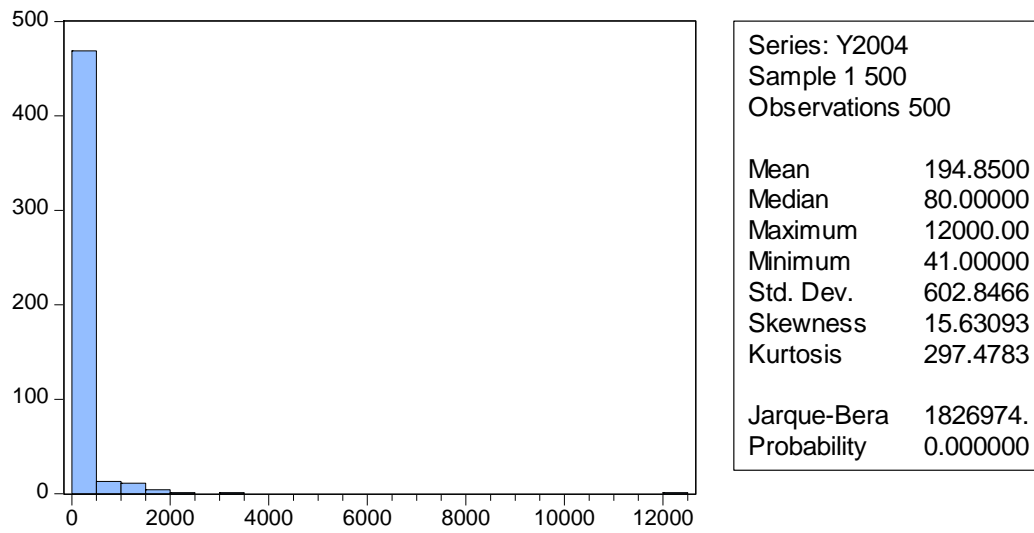
Figure 1: A typical graph of the data
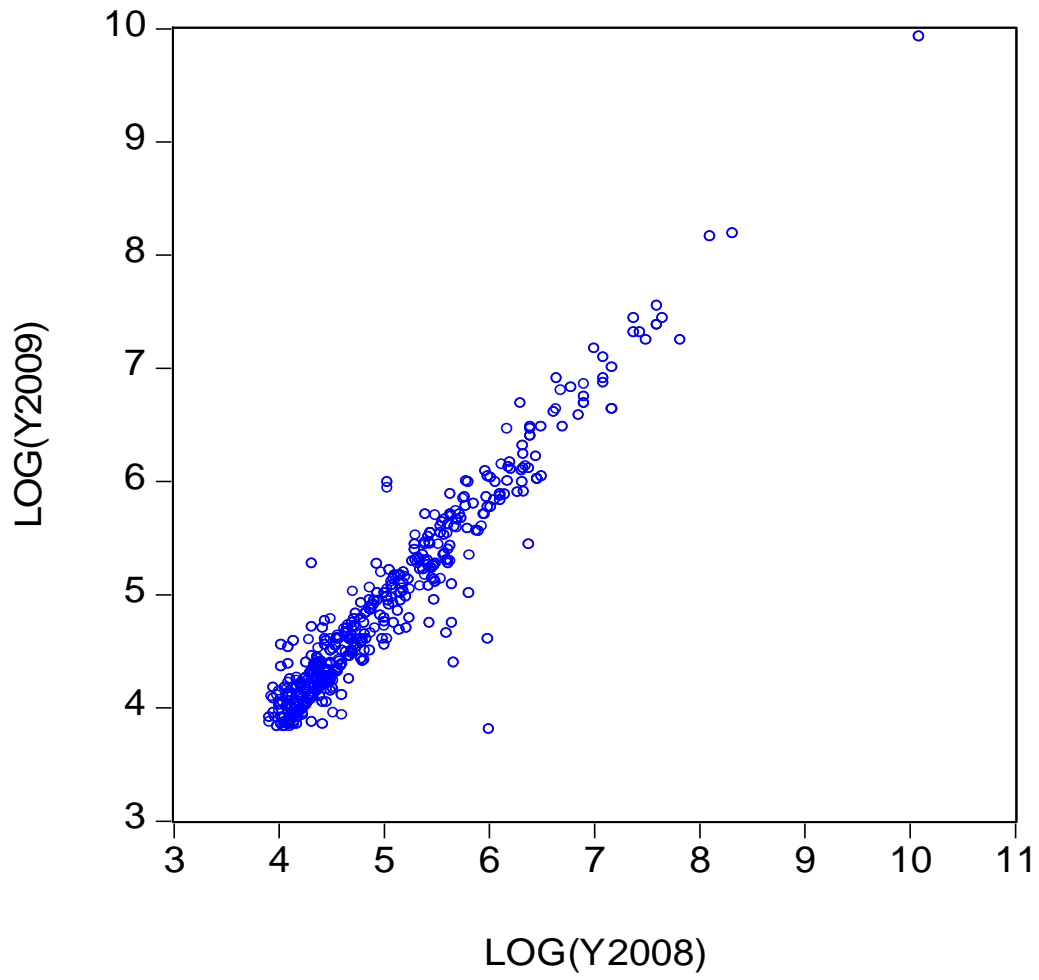
Figure 2: Annual changes, visualized
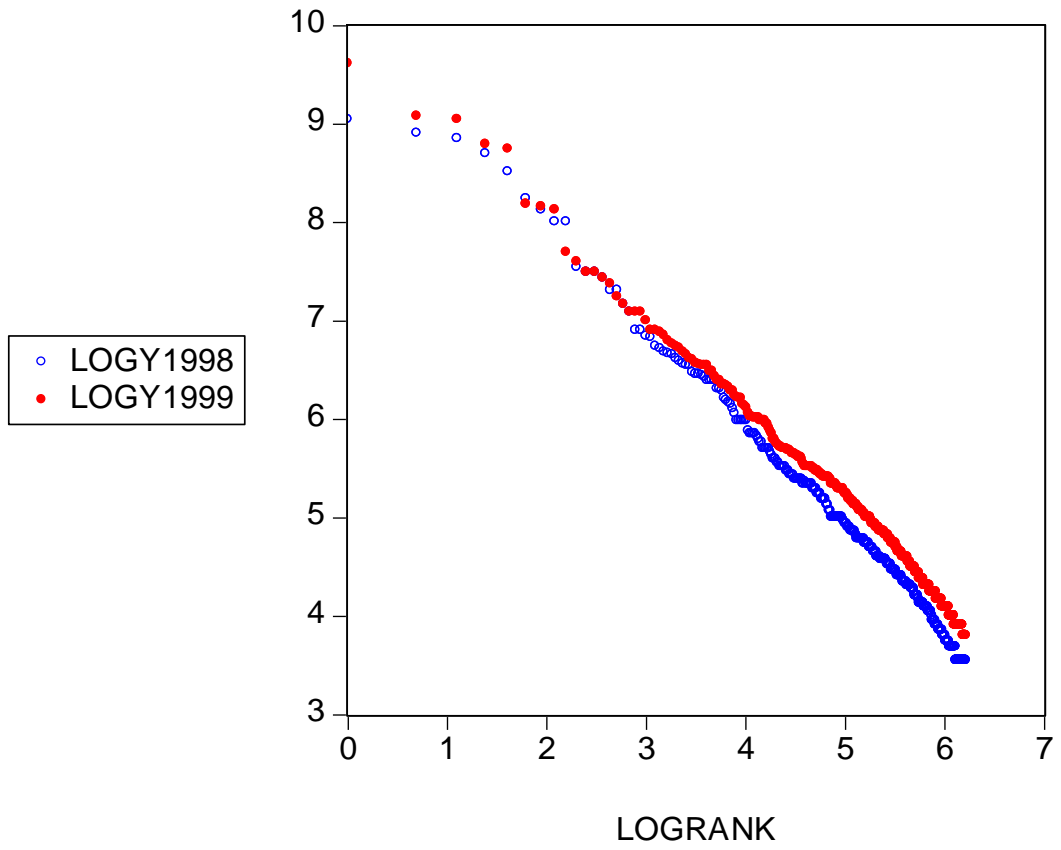
Figure 3a: The power law visualized, 1998-1999
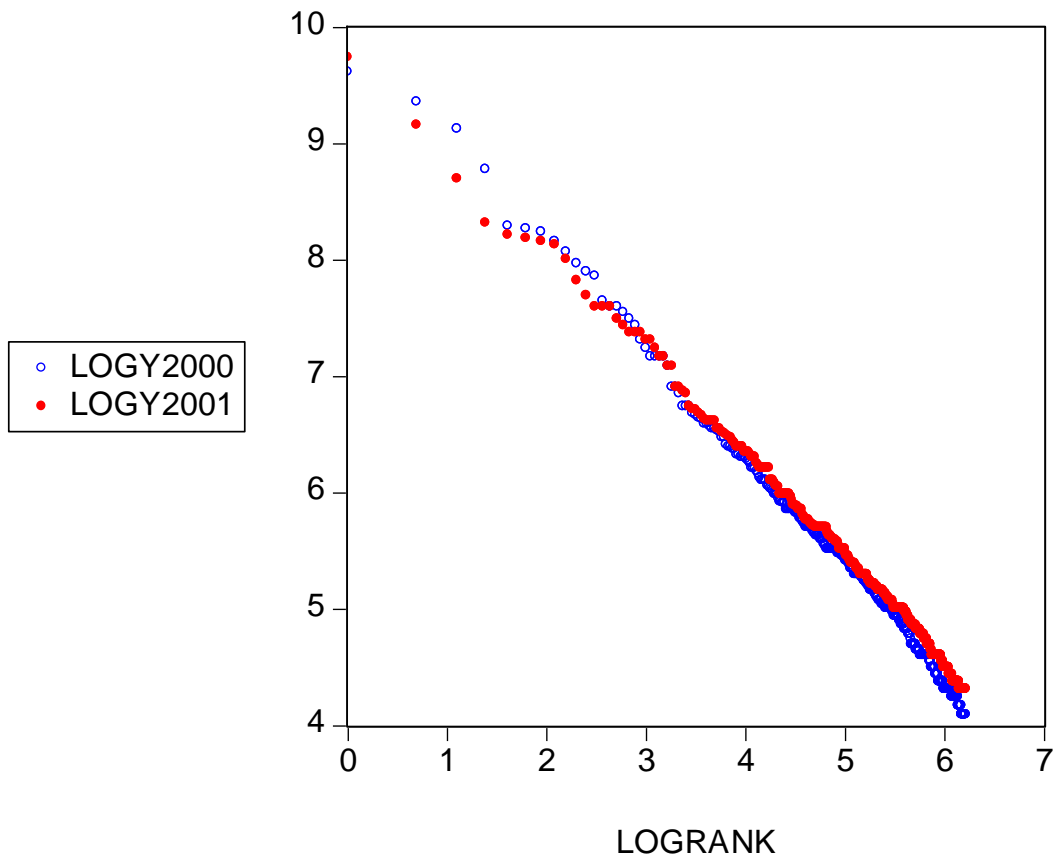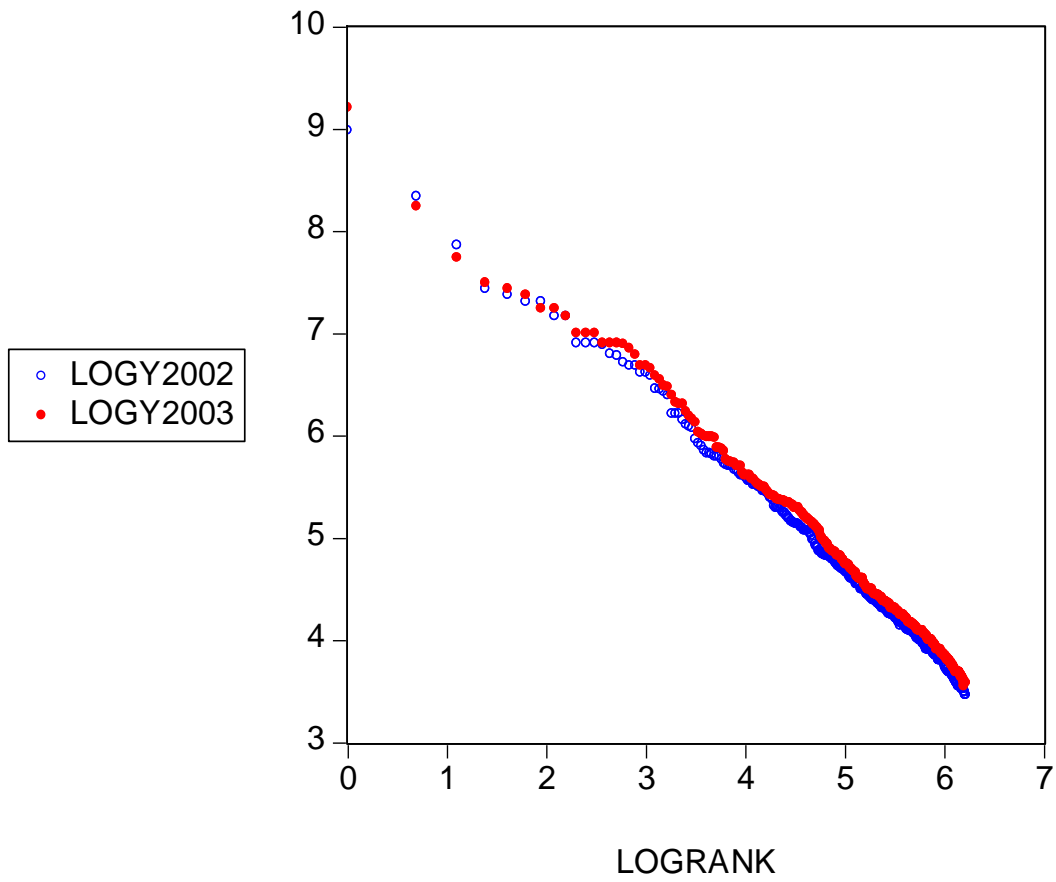
Figure 3b: The power law visualized, 2000-2001

Figure 3c: The power law visualized, 2002-2003
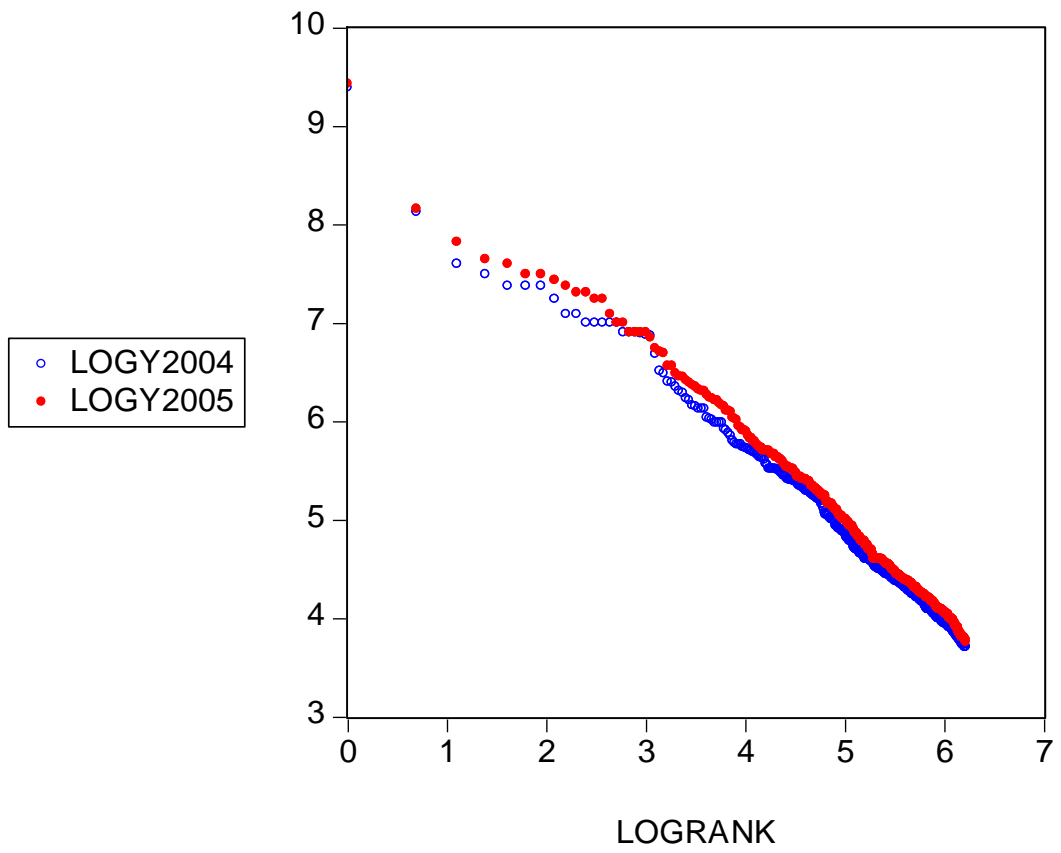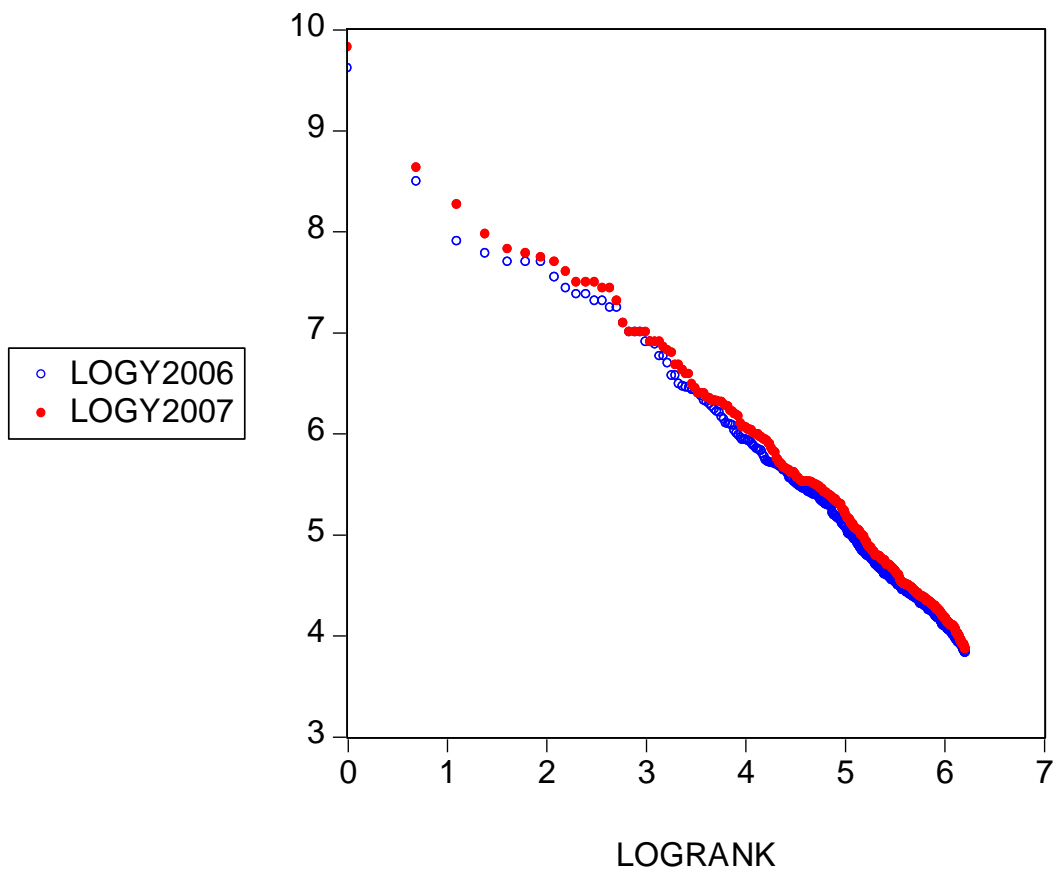
Figure 3d: The power law visualized, 2004-2005
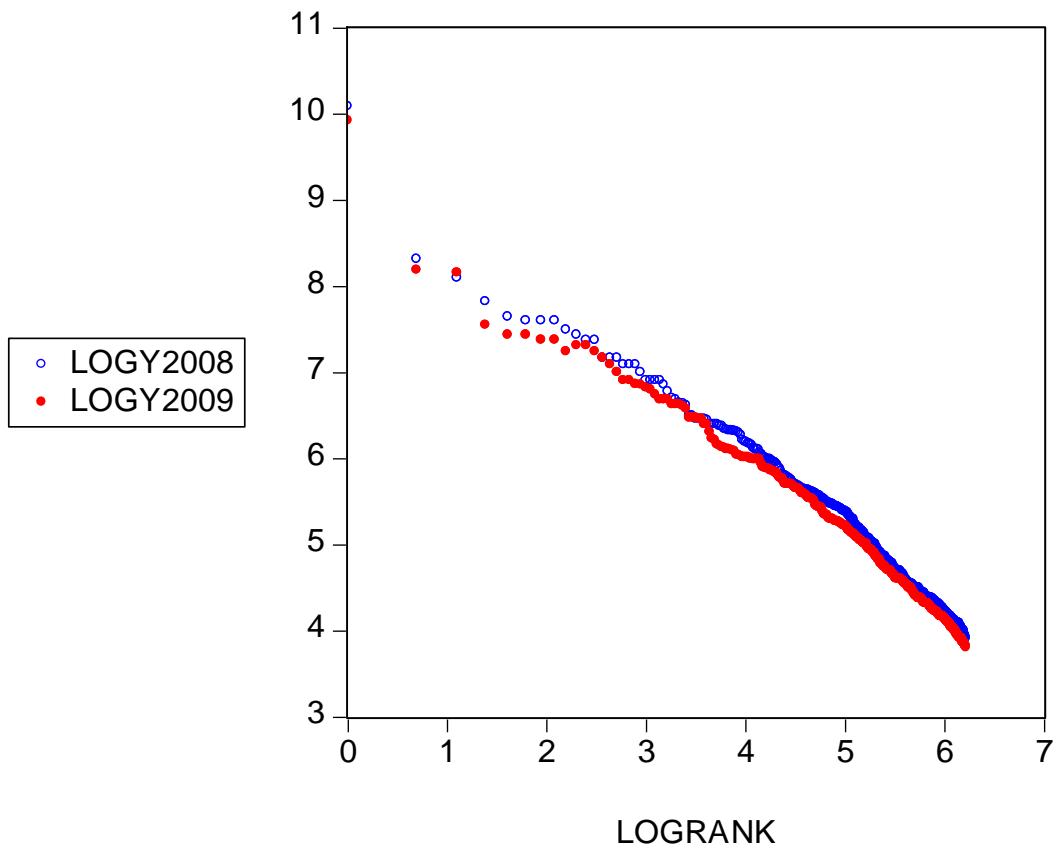
Figure 3e: The power law visualized, 2006-2007
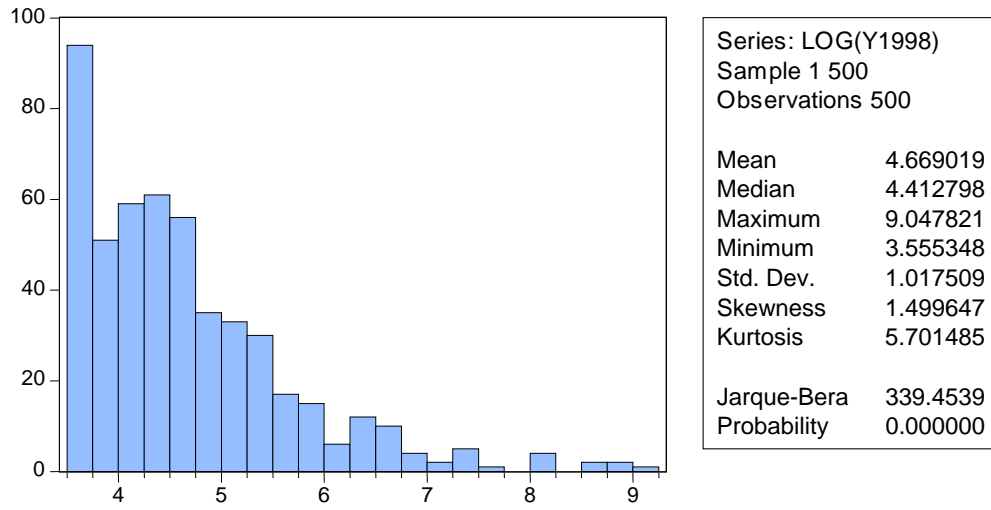
Figure 3f: The power law visualized, 2008-2009

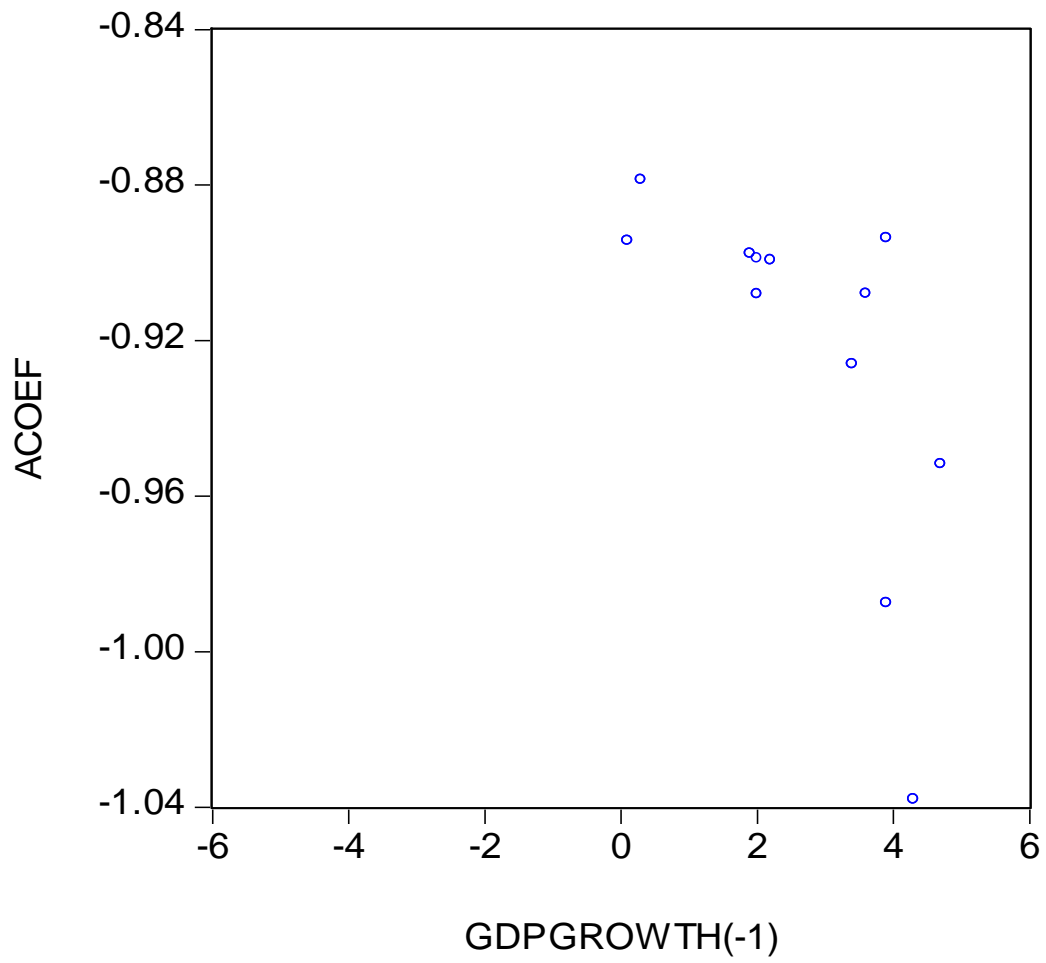|	Series: LOG(Y1998)
	Sample 1 500
	Observations 500
Mean	4.669019
Median	4.412798
Maximum	9.047821
Minimum	3.555348
Std. Dev.	1.017509
Skewness	1.499647
Kurtosis	5.701485
Jarque-Bera	339.4539
Probability	0.000000

Figure 4: the Logs of wealth, a typical graph

Figure 5: Inequality versus one-year lagged growth

References

Barro, R.J. (2000), Inequality and growth in a panel of countries, *Journal of Economic Growth*, 5, 5-32.

Castaldi, C. and M. Milakovic (2007), Turnover Activity in Wealth Portfolios, *Journal of Economic Behavior & Organization*, 63, 537-552

Davies, J.B. and A.F. Shorrocks (2000), The Distribution of Wealth, in A.B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, Amsterdam: Elsevier, 605-675.

Levy, M. and S. Solomon (1997), New Evidence for the Power Law Distribution of Wealth, *Physica A*, 242, 90-94.

Neumayer, E. (2004), The Super-Rich in Global Perspective: A Quantitative Analysis of the Forbes List of Billionaires, *Applied Economics Letters*, 11, 793-796.

Newman, M.E.J. (2005), Power Laws, Pareto Distributions and Zipf's Law, *Contemporary Physics*, 46, 323-351.

Persky, J.J. (1992), Pareto's Law, *Journal of Economic Perspectives*, 6, 181-192.