A Multivariate Statistical Analysis of the Free World

David Friedenberg
Miami University
Oxford OH

Shenek Heyward
Francis Marion University
Florence, SC

Abstract

Is a democracy more than just competitive multiparty elections in which all participants have a legitimate chance of attaining power? Using such statistical analyses processes such as discriminant analysis and factor analysis, we hope to determine a rule for classifying countries from a sample into one of two groups, democratic or non-democratic. We also hope to reduce our data from 11 variables to a smaller set of underlying factors that can be used to explain the dynamics surrounding each country.

> *We must seize the vision of a free and diverse world—*
> *and shape our policies to speed progress toward a flexible*
> *world order.*
>
> –President John F. Kennedy

Introduction

The Cold War pitted, in a head to head confrontation, the democracies of the world against communism. While the Cold War may be over, the concept of spreading democracy and freedom to the world is still forefront in America's foreign policy. Whether in war torn Afghanistan or developing nations in Africa, the United States actively encourages democracy. But, in this regard, our leaders face a number of questions. What characteristics do democratic nations share? Can we classify nations based on certain characteristics? And how should we spend our money and exert our influence as to get the best results?

This paper seeks to explore and analyze data about countries in the hope that we can statistically address the questions posed above. The first technique we will employ is *factor analysis*. By using factor analysis we hope to find underlying factors in our model that we can label and measure. These factors will reduce the dimensionality of our data set, thereby making it simpler and easier to comprehend. The second technique we will use is *discriminant analysis*. Using this technique we will develop a model to predict whether a country can be classified as democratic or not, based on certain measurements.

The data we used came from the CIA World Factbook [1], a database that catalogs information on the U.S. and foreign nations. We looked for numeric data that would accurately, but concisely describe a country. We decided on 11 variables that described a country's size, economy, military, and education system. The 11 variables are as follows: area, land use (percent arable land), population, population growth rate, life expectancy, literacy, gross domestic product (GDP), number of televisions per capita, number of airports per capita, military expenditure, and percent of gross domestic product spent on military.

Assessment of Normality

Our next concern was the multivariate normality of the population from which the sample data was obtained. Both factor analysis and discriminant analysis assume multivariate normality of the data, so we checked our data for normality. To check for normality we constructed Q-Q plots, a plot of ordered points such that the ordered point ($x_{(j)}$, $q_{(j)}$) is an ordering of the ordered statistics and the ordered normal quantiles. We wanted the Q-Q plots of the points to have a linear relationship indicating a high correlation. To determine if a variable was normally distributed, the null hypothesis $H_0: \rho=1$ versus $H_1: \rho<1$ had to be tested where ρ is the population correlation coefficient. Thus, we calculated the correlation between the ordered data, $x_{(j)}$ and the ordered normal quantiles, $q_{(j)}$, at $\alpha=0.01$. If the variable had a perfect correlation of 1 with the ordered normal quantiles, then it was determined that the variable was normally distributed. In practice $r_{(Q)}$, the sample correlation coefficient was calculated where $r_{(Q)} = \dfrac{\mathrm{cov}(x_{(j)}, q_{(j)})}{(\sqrt{\mathrm{var}\, x_{(j)}})(\sqrt{\mathrm{var}\, q_{(j)}})}$ (Appendix 1.C). The null hypothesis $H_0: \rho=1$ was rejected if $r_{(Q)} <$ the critical value. For our data, the critical value was 0.9538 at a significance level $\alpha=0.01$ [4]. The population growth rate, life expectancy, and literacy rate were all normally distributed at a significance level of $\alpha=0.01$. Even though the remaining eight variables failed the test of normality, due to the robust nature of the multivariate procedures, we were allowed to proceed without making transformations on these eight variables.

Factor Analysis Theory

Factor analysis is a multivariate statistical technique that we will use to reduce our data from 11 variables to a smaller set of underlying factors. These factors will be simpler to analyze and in general, can also be used as inputs for discriminant analysis.

The Factor Analysis model is used on multivariate, normally distributed variables, $\underline{x} \sim N_p(\underline{\mu}, \Sigma)$ where $\underline{\mu}$ and Σ are given by $\underline{\mu} = [\mu_1 \ \cdots \ \mu_p]'$ and

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ & \sigma_{22} & \cdots & \sigma_{2p} \\ & & \ddots & \vdots \\ & & & \sigma_{pp} \end{bmatrix}$$ respectively.

The factor analysis model is based on m underlying factors, where $m<p$. The factors are assumed common to each variable x_i, while there are p specific factors η_i, each of which applies only to the corresponding x_i . . The m factor model is:

$$x_i = \lambda_{i1}F_1 + \lambda_{i2}F_2 + \ldots + \lambda_{im}F_m + \eta_i \quad \text{for } i = 1,2,\ldots p.$$

$$\text{or in matrix notation } \underline{x} = \Lambda\underline{F} + \underline{\eta} \; .$$

Each component λ_{ik} of the matrix Λ is called a *factor loading*, specifically the loading of the i^{th} variable on the k^{th} factor. A factor loading, $\lambda_{ik,}$ measures the contribution and importance of the common factor F_k to the response variable x_i, and $\sum_{k=1}^{m}\lambda_{ik}^2$ measures variance in x_i accounted for by the common factors.

This m factor model is based on the following three assumptions:

1. F_k ~ iid N(0,1)
2. $\eta_i \sim N(0, \Psi_i)$ for $i = 1,2\ldots p$ where Ψ is a diagonal matrix
3. $F_k \perp \eta_i$ for all (K, i)

It can be shown that because of the m-factor model assumptions that $\text{cov}(\underline{x}) = \Sigma = \Lambda\Lambda' + \Psi$.

In theory, we decide on an initial number of common factors, m, and using the above equation we obtain the matrices Λ and Ψ. However in reality, we do not know Σ so we use the sample variance-covariance matrix, S, to obtain estimates. Maximum Likelihood Estimation is used to obtain the factor loadings. The factor loadings we obtain are not unique. Therefore, Λ is not the only loading matrix to solve the equation $\Sigma = \Lambda\Lambda' + \Psi$, and hence $\hat{\Lambda}$.

Once solutions are found, they can be tested to see if they satisfy the hypothesis H_0: $\Sigma = \Lambda\Lambda' + \Psi$. The likelihood ratio test (LRT) is used to check the suitability of the chosen number of factors, m. The testing is repeated until the minimum value of m is obtained for which we accept H_0 at a chosen level of significance, α. Once m is determined, the excised factors must be interpreted and appropriately labeled. Thus the original data is reduced from p variables to m factors, which are ready to use for further analysis.

Factor Analysis Application

In Factor Analysis we want the variables x_1 , x_2 , … x_p to be dependent. Thus we test for complete independence of the variables by computing the correlation matrix P. We test H_0: P=I versus H_1:P$\neq$I. We use a likelihood ratio test to test H_0 at α=0.05. The null hypothesis H_0 : P=I is rejected if the test statistic > the critical value. For our data the critical value was $\chi^2_{55,\, 0.05}$ =73. 3115. The test statistic was calculated using the formula $-(n - 1 - \frac{2p+5}{6})\ln|P|$. Our test statistic was 301.2183, and it follows that the P-value was 0 which is a very strong indication that H_0 should be rejected. Thus, we rejected the null hypothesis because the test statistic was greater than the critical value. This test tells us that the variables are dependent so we proceed to create a Factor model. We used Minitab to create a 3-factor model. We then tested the adequacy using a Likelihood Ratio Test. With this test we accept the *m*-factor model if the test statistic,

$$[n\text{-}1\text{-} \frac{1}{6}(2p\text{+}5) \text{ - } \frac{2}{3}m] \ln \frac{|\hat{\Psi} + \hat{\Lambda}\hat{\Lambda}'|}{|R|} < \chi^2_{\upsilon,\alpha}$$

, where υ is the degrees of freedom and α is the significance level. We set α=0.05. The following table shows degrees of freedom, test statistics, and critical values at α=0.05 for the *m* factor models we tested.

Table 1.1: *m* factor adequacy test

m Factor model	Degrees of Freedom	Test Statistic	Critical Value	P-value
1 Factor model	44	129.2393	60.4809	0
2 Factor model	34	69.9622	48.6024	0.003
3 Factor model	25	32.0437	37.6525	0.1568

Comparing the test statistic for the *m*-factor model to its corresponding critical value, the first *m* for which the test statistic < the critical value is for *m*=3. Hence, $H_0: \Sigma = \Lambda\Lambda' + \Psi$ is accepted for *m* = 3 at α = 0.05. Therefore, we determined that the 3-factor model was adequate. The following table shows the factor loadings on all of the variables for the 3 factor model.

Table 1.2 : 3 Factor Model

Variable	Factor 1 Loadings	Factor 2 Loadings	Factor 3 Loadings
Area	0.602	0.165	-0.011
Land Use	0.510	-0.259	0.188
Population	0.989	0.088	0.013
Population Growth Rate	-0.063	0.611	-0.358
Life Expectancy	0.079	-0.708	-0.260
Literacy Rate	-0.139	-0.878	0.050
GDP	1.000	0.000	0.000
Televisions per capita	-0.091	-0.654	-0.434
Airports per capita	-0.157	-0.028	0.014
Military Expenditure	0.774	-0.158	-0.413
% GDP spent on Military	-0.019	0.011	-0.993
Variance	3.2620	2.2006	1.5205
% Variance	0.297	0.200	0.144

Magnitudes and signs of factor loadings are used to determine factor labels. The first factor had very high positive loadings on GDP, population, and military expenditures. Therefore we called this factor the *Wealth and Power Factor*. This factor accounts for 29.7% of the variation in the data. The second factor had a high positive loading on population growth rate and had high negative loadings on literacy rate, life expectancy, and televisions per capita. Thus, we decided to call factor two the *Lack of Development Factor*. The *Lack of Development/Development* Factor accounts for 20% of the variance in the data. The final factor we called the *Low Military Priority Factor* as it contained a high negative loading on % GDP spent on military. This factor accounted for 14.4% of the variance in the data. Our 3- factor model was adequate and accounted for approximately 64% of the variation in the data. In Social Sciences a model that accounts for 60% of the variation or more is acceptable. Therefore, in addition to the acceptance of H_0: $\Sigma = \Lambda\Lambda' + \Psi$ for $m = 3$, we believed our 3-factor model was a good model, which reduced the dimensionality of our data set from 11 to 3.

Theory of Discriminant Analysis

In many aspects of life it is necessary to classify objects. Creditors need to know whether someone is a good or a poor credit risk. A college admissions board needs to determine if an applicant is likely to be successful or unsuccessful. A surgeon has to determine whether a patient is a good or a poor candidate for an organ transplant. Discriminant analysis, which is sometimes called classification analysis, is very useful in determining a rule for classifying people or objects from a sample into one of several categories.

If there are two *p*-variate populations where $\Pi_1 \equiv N_p(\underline{\mu}_1, \Sigma)$ and $\Pi_2 \equiv N_p(\underline{\mu}_2, \Sigma)$, we can obtain a classification rule to classify $\underline{x}$, an unclassified sample element. The assumption is made that $\underline{x} \sim N_p(\underline{\mu}, \Sigma)$ and
$\underline{x} = [x_1 \quad x_2 \quad . \quad . \quad x_p .]'$. In practice, the mean and covariance of each population will be unknown so samples from Π_1 and Π_2 are used to create the classification rule. These samples are called training samples and are of sizes n_1 and n_2. The training samples in addition to being used for creation of the classification rule, also are used to obtain estimates of the mean and covariance matrices of each population. The estimates of $\underline{\mu}_1$, $\underline{\mu}_2$, Σ_1, and Σ_2 are $\bar{\underline{x}}_1$, $\bar{\underline{x}}_2$, S_1, and S_2, respectively.

It has been shown that if the covariance matrices of Π_1 and Π_2 are equal, then the estimates of $\underline{\mu}_1$, $\underline{\mu}_2$, and Σ are $\bar{\underline{x}}_1$, $\bar{\underline{x}}_2$, and $S_p = \dfrac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2}$, where S_p is a pooled estimate from Π_1 and Π_2 of Σ. The expected value of each of the estimates S_p, S_1, and S_2 is Σ. The estimates S_p, S_1, and S_2 are all unbiased estimates of Σ.

We want to classify $\underline{x}$ into Π_1 or Π_2 on the basis of a linear discriminant function $\underline{a}'\underline{x}$ where $\hat{\underline{a}} = S_p^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)$. Using the linear discriminant function, we classify $\underline{x}$ into Π_1 or Π_2 according to the following rule:

$$\text{classify } \underline{x} \text{ into } \Pi_1 \text{ if } \hat{\underline{a}}'\underline{x} > \hat{h}$$

$$\text{or classify } \underline{x} \text{ into } \Pi_2 \text{ if } \hat{\underline{a}}'\underline{x} \leq \hat{h}$$

$$\text{where } \hat{\underline{a}} = S_p^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) \text{ and } \hat{h} = \frac{1}{2}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)'S_p^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2).$$

Once we have made our classification rule, it is important to find the *apparent error rate* (AER), which measures how well our model is classifying the training sample. The apparent error rate is the percentage misclassified from the training samples using the classification rule. The apparent error rate can be found by subtracting the percentage of correctly classified items from the training sample from 100%. The total probability of misclassification (TPM) is the rate of misclassification of sample elements using the sample classification function. The distance between the two populations must be found in order to determine the TPM. This distance is denoted by Δp^2 where

Δp^2 = (Mahalanobis distance between Π_1 and $\Pi_2)^2$ = $\underline{a}'\Sigma\underline{a}$ = $(\underline{\mu}_1 - \underline{\mu}_2)'\,\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$.

After determining this distance, the TPM can be found where the TPM= $2\Phi(-\frac{1}{2}\Delta p)$ and $\Sigma_1=\Sigma_2=\Sigma$. For this purpose, $\Phi(x)$ is the cumulative distributive function of a standard normal distribution with mean 0 and variance 1. In practice, we find a TPM which is approximately equal to

$$\hat{\alpha} = 2\Phi(-\frac{1}{2}\hat{\Delta}_p), \text{ and } \hat{\Delta}_p^2 = (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)'S_p^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2).$$

The classification rule assigns $\underline{x}$ to Π_1 if $\hat{a}'\underline{x} > \hat{h}$, which is equivalent to assigning $\underline{x}$ to Π_1 if the Mahalanobis distance (D_1^2) of $\underline{x}$ from Π_1 is less than the Mahalanobis distance (D_2^2) of $\underline{x}$ from Π_2. Thus if $(\underline{x} - \underline{\mu}_1)'\Sigma^{-1}(\underline{x} - \underline{\mu}_1) < (\underline{x} - \underline{\mu}_2)'\,\Sigma^{-1}(\underline{x} - \underline{\mu}_2)$ then $\underline{x}$ is classified into Π_1 (and similarly for classifying $\underline{x}$ into Π_2). Theory shows that this procedure is optimal because it minimizes the TPM. In general, using the minimum distance classification rule for the *k* population case, classify $\underline{x}$ into Π_j if $D_j^2 = \text{Min}(D_1^2, D_2^2, \ldots, D_k^2)$. The linear discriminant function is equivalent to the minimum distance classification.

In order to use the linear discriminant function rule to classify $\underline{x}$ into one of the two populations, it is assumed that $\Sigma_1 = \Sigma_2$.

Discriminant Application

For discriminant analysis, we needed to classify the countries in our training sample as either democratic or not democratic. For help on this, we consulted Dr. Ryan Barilleaux, the Chair of the Political Science Department at Miami University. He showed us the Freedom House website and a list of all the world's democratic countries [5]. We used this list to separate our training samples into $\Pi_{1,}$ democratic countries, and $\Pi_{2,}$ non-democratic countries. Our training sample was based on a random selection of 36 countries from a list of 191 countries that were listed in the World Factbook. From our original random sample of 50 entries in the World Factbook we eliminated fourteen because either they were regions, not countries (eg. Antartica), or because there was incomplete data for a country.

.

To justify using a linear discriminant function, we needed to show $\Sigma_1 = \Sigma_2$. We tested H_0: $\Sigma_1 = \Sigma_2$ versus H_1 : $\Sigma_1 \neq \Sigma_2$ using the likelihood ratio test. Our test statistic was $\frac{m}{c}$, and we tested the test statistic against $\chi^2_{66,\ 0.05} = 85.9649$. We calculated $\frac{m}{c}$ for our data set to be 183.25 where

$$m = \{\sum_{i=1}^{2}(n_i - 1)\}\{\ln |S_p|\} - \sum_{i=1}^{2}(n_i - 1)\ln |S_i| \text{ and}$$

$$\frac{1}{c} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)}\{(\frac{1}{(n_1 - 1)} + \frac{1}{(n_2 - 1)}) - \frac{1}{n_1 + n_2 - 2)}\},$$ and the P-value was 0 which is a strong indication that H_0 should be rejected. Since $\frac{m}{c}$ is greater than our critical value we rejected the null hypothesis. Since the linear discriminant model assumes $\Sigma_1 = \Sigma_2$, our rejection of $\Sigma_1 = \Sigma_2$ tells us that we should not use a linear discriminant model, but instead should use a quadratic model. The following table shows how the Quadratic model classified the training sample:

Table 1.3: Quadratic Model

	True Group Democracy	True Group Non-Democracy
Classified into Group		
Democracy	17	0
Non-Democracy	3	16
Total N	20	16
N Correct	17	16
Proportion	0.850	1.000
N=36	N correct = 33	Proportion Correct = 0.917

We noticed that our model had an apparent error rate of 1-0.917 = 0.083. This said that our model did a good job of classifying the countries in our training sample. However, the main purpose of creating this discriminant model was to classify countries that were not in the training sample and typically unknown. We randomly selected 23 other countries to classify using this model. Since we had a list of all the democratic countries in the world, we were able to check how well our discriminant model was classifying these "unknown" countries. Our model only classified 13 of the 23 countries correctly. Obviously, this was disappointing; thus we attempted to create a better model.

When we reviewed our original model we noticed that there were high correlations between Population, Number of Televisions, and Number of Airports. We hypothesized that if we divided the Number of Televisions and Number of Airports by the corresponding Population that our model would improve.

With the aforementioned changes, we tested $H_0 : \Sigma_1 = \Sigma_2$ versus $H_1 : \Sigma_1 \neq \Sigma_2.$ Similar to the previous test of $\Sigma_1 = \Sigma_2$, our test statistic was greater than the critical value; thus, we rejected the null hypothesis that $\Sigma_1 = \Sigma_2$. Even though $\Sigma_1 \neq \Sigma_2$, we were still able to use the linear discriminant function to classify the countries in our training sample because of the robustness of the tests used in discriminant analysis. The following table shows how the linear discriminant model classified the training sample:

Table 1.4: Linear Discriminant Model of Transformed Data

	True Group Democracy	True Group Non-Democracy
Classified into Group		
Democracy	18	3
Non-Democracy	2	13
Total N	20	16
N Correct	18	13
Proportion	0.900	0.812
N=36	N correct = 31	Proportion Correct = 0.861

Again we also used the quadratic model to classify our data because this model can be used whether or not Σ_1 and Σ_2 are equal. The following table shows how the Quadratic model classified the training sample:

Table 1.5: Quadratic Model of Transformed Data

	True Group Democracy	True Group Non-Democracy
Classified into Group		
Democracy	18	0
Non-Democracy	2	16
Total N	20	16
N Correct	18	16
Proportion	0.900	1.000
N=36	N correct = 34	Proportion Correct = 0.944

We decided that the best model for classifying both our training sample and the 23 randomly selected countries not in our training sample was obtained using the quadratic classification model on the transformed data set. The apparent error rate for the quadratic model of the transformed data was 0.056 which was smaller than 0.139, the AER of the linear transformed model and smaller than 0.083, the AER of the quadratic model of the original data. The quadratic model of the transformed data also performed as well as or better than the other two models when classifying a random sample of "unknown" countries. Both the quadratic and linear models of the transformed data classified 16 out of 23 of the "unknown" countries correctly, but the quadratic model of the original data only classified 13 of the 23 "unknown" countries correctly. The following table shows how the quadratic model of the transformed data classified our "unknown" sample of 23 countries.

Table 1.6: Classification of “Unknown” Sample by Quadratic Model on Transformed Data

Observation	Country	Predicted Group	Squared Distance to π_1	Squared Distance to π_2
1	Angola	N	1083.918	393.001
2	Bahrain	N	168.554	130.778
3	Belize	D	133.105	965.429
4	Burkina Faso	D	123.307	136.874
5	Burundi	N	183.352	145.118
6	Djibouti	N	152.316	111.052
7	Finland	D	145.623	157.311
8	Ghana	N	107.263	103.244
9	Iran	D	171.337	400.736
10	Latvia	D	112.999	186.199
11	Morocco	N	142.185	110.326
12	Nepal	N	129.822	111.322
13	Niger	D	188.560	193.679
14	Peru	N	136.757	128.654
15	Senegal	N	126.237	111.746
16	Seychelles	D	166.150	626.797
17	Sierra Leone	N	120.851	117.213
18	Singapore	D	132.850	353.776
19	South Africa	D	176.787	399.620
20	Sri Lanka	N	143.373	106.551
21	Thailand	D	197.382	711.882
22	Turkmenistan	N	124.212	111.523
23	United States	D	221368.629	1180148.242

This quadratic model correctly classified 16 of the 23 countries that were not in the training sample. Thus our model correctly classified countries not in our training sample approximately 70% of the time.

Conclusion

We were satisfied with our factor model that was obtained in the factor analysis because it greatly reduced the dimensionality of our data set from 11 variables to only 3. However, the 3-factor model still accounted for an acceptable level of the variance explained by all 11 variables. Though our discriminant analysis model worked well classifying the countries in our training sample, we were not completely satisfied with the percentage of countries it classified correctly that were not in our training sample. We believe that our discriminant analysis model may be improved by adding such variables as unemployment rate, religion, and GDP per capita to our model. We also concluded that in order for our model to be effective year to year, that the model will have to be periodically updated because some variables such as area will tend to remain constant, but variables such as GDP and military expenditures will vary in the future. We have demonstrated how multivariate statistics can be used to analyze countries that are similar and also different from our own. We hope we have shown how multifaceted the word democracy is and come a bit closer in understanding the dynamics that surround it for with understanding there can be improvement. Because how can we hope to help those who have never tasted of freedom, if we do not truly understand what democracy is ourselves?

Mahatma Gandhi said, "I understand democracy as something that gives the weak the same chance as the strong."

Acknowledgments

We would like to thank Dr. Vasant Waikar for guiding us through this paper, teaching us the statistics, and helping us with the writing. We would also like to thank Lindsay Kellam for her infinite patience and help in everything we did or tried to do. We would also like to thank Dr. Ryan Barilleaux of Miami University for his expertise. Finally we would like to thank all the students, graduate assistants, and professors at SUMSRI, without your support this paper would not have been possible.

Bibliography

[1] CIA World Factbook. (2001). Retrieved June 25, 2002 from http://www.odci.gov/cia/publications/factbook/

[2] W.R. Dillon and M. Goldstein, Multivariate Analysis (John Wiley & Sons, New York, 1984).

[3] D. Johnson, Applied Multivariate Methods for Data Analysis (Duxbury, Pacific Grove, 1998).

[4] R.A. Johnson and D.W. Wichern, Applied Multivariate Statistical Analysis 4th edition (Prentice Hall, Upper Saddle River NJ, 1998).

[5] Table of Electoral Democracies. (2000). Retrieved July 2, 2002 from http://www.freedomhouse.org/research/freeworld/2000/table7

Appendix 1.A Data for Democratic Countries

Country	Area	Land use	Population	Pop Growth	Life expectancy	Literacy	GDP
Albania	28748	0.21	3510484	0.0088	71.83	0.93	1.05E+10
Benin	112620	0.13	6590782	0.0297	49.94	0.375	6.60E+09
Botswana	600370	0.01	1586119	0.0047	37.13	0.698	1.04E+10
C.A.R.	622984	0.03	3576884	0.0185	43.8	0.6	6.10E+09
Georgia	69700	0.09	4989285	-0.0059	64.57	0.99	2.28E+10
Greece	131940	0.19	10623835	0.0021	78.59	0.95	1.82E+11
Guatemala	108890	0.12	12974361	0.026	66.51	0.636	4.62E+10
Honduras	112090	0.15	6406052	0.0243	69.35	0.727	1.70E+10
Hungary	93030	0.51	10106017	-0.0032	71.63	0.99	1.14E+11
India	3287590	0.56	1029991145	0.0155	62.86	0.52	2.20E+12
Indonesia	1919440	0.1	228437870	0.016	68.27	0.838	6.54E+11
Israel	20770	0.17	5938093	0.0158	78.71	0.95	1.10E+11
Lithuania	65200	0.39	3610535	-0.0027	69.25	0.98	2.64E+10
Madagascar	587040	0.04	15982563	0.0302	55.35	0.8	1.23E+10
Namibia	825418	0.01	1797677	0.0138	40.62	0.38	7.60E+09
Panama	78200	0.07	2845647	0.013	75.68	0.908	1.66E+10
Paraguay	406750	0.06	5734139	0.026	73.92	0.921	2.62E+10
Sao Tome	1001	0.02	165034	0.0318	65.59	0.73	1.78E+08
Uruguay	176220	0.07	3360105	0.0078	75.44	0.973	3.10E+10
Venezuala	912050	0.04	23916810	0.0156	73.31	0.911	1.46E+11

Appendix 1.A continued

Country	Televisions per capita	Airports per capita	Military expenditures	Military exp % GDP
Albania	0.115369	0.0000031	4.20E+07	0.015
Benin	0.009104	0.0000008	2.70E+07	0.012
Botswana	0.019545	0.000058	6.10E+07	0.012
C.A.R.	0.005032	0.0000145	2.90E+07	0.022
Georgia	0.515104	0.0000062	2.30E+07	0.0059
Greece	0.239085	0.0000076	6.12E+09	0.0491
Guatemala	0.10197	0.0000368	1.20E+08	0.006
Honduras	0.088978	0.0000186	3.50E+07	0.006
Hungary	0.437363	0.0000043	8.22E+08	0.016
India	0.061166	0.0000003	1.30E+10	0.025
Indonesia	0.060191	0.000002	1.00E+09	0.013
Israel	0.284603	0.0000093	8.70E+09	0.094
Lithuania	0.470844	0.0000199	1.81E+08	0.0166
Madagascar	0.020335	0.0000081	2.90E+07	0.01
Namibia	0.033376	0.0000729	1.04E+08	0.026
Panama	0.179221	0.0000376	1.28E+08	0.013
Paraguay	0.17265	0.0001596	1.25E+08	0.014
Sao Tome	0.139365	0.0000121	1.00E+06	0.015
Uruguay	0.232731	0.000019	1.72E+08	0.009

Appendix 1.B Data for Non-Democratic Nations

Countries	Area	Land use	Population	Pop Growth rate	Life expectancy	Literacy	GDP
Algeria	2381740	0.03	31736053	0.0171	69.95	0.616	1.71E+11
Azerbaijan	86600	0.18	7771092	0.0032	62.96	0.97	2.35E+10
Belarus	207600	0.29	10350194	-0.0015	68.14	0.98	7.88E+10
Burma	678500	0.15	41994678	0.006	55.16	0.831	6.37E+10
D.R Congo	2345410	0.03	53624718	0.031	48.94	0.773	3.10E+10
Equatorial Guinea	28051	0.05	486060	0.0246	53.95	0.785	9.60E+08
Ethiopia	1127127	0.12	65891874	0.027	44.68	0.355	3.92E+10
Gabon	267667	0.01	1221175	0.0102	49.59	0.632	7.70E+09
Kazakhstan	2717300	0.12	16731303	0.0003	63.29	0.98	8.56E+10
Kuwait	17820	0	2041961	0.0338	76.27	0.786	2.93E+10
Libya	1759540	0.01	5240599	0.0242	75.65	0.762	4.54E+10
Oman	212460	0	2622198	0.0343	72.04	0.8	1.96E+10
Togo	56785	0.38	5153088	0.026	54.35	0.517	7.30E+09
Tunisia	163610	0.19	9705102	0.0115	73.92	0.667	6.28E+10
Uganda	236040	0.25	23985712	0.0293	43.37	0.618	2.62E+10
Yemen	527970	0.03	18078035	0.0338	60.21	0.38	1.44E+10

Appendix 1.B continued

Countries	Televisions per capita	Airports per capita	Military expenditures	Military exp % GDP
Algeria	0.097681	0.0000043	1.87E+09	0.041
Azerbaijan	0.021876	0.0000067	1.21E+08	0.026
Belarus	0.243474	0.0000131	1.56E+08	0.012
Burma	0.00762	0.0000019	3.90E+07	0.021
D.R Congo	0.120803	0.0000043	2.50E+08	0.046
Equatorial Guinea	0.008229	0.0000062	3.00E+06	0.006
Ethiopia	0.004856	0.0000013	1.38E+08	0.025
Gabon	0.05159	0.0000483	9.10E+07	0.016
Kazakhstan	0.231901	0.0000268	3.22E+08	0.015
Kuwait	0.42851	0.0000039	1.90E+09	0.087
Libya	0.139297	0.000026	1.30E+09	0.039
Oman	0.610175	0.0000545	2.40E+09	0.13
Togo	0.014166	0.0000017	2.70E+07	0.02
Tunisia	0.094796	0.0000033	3.56E+08	0.015
Uganda	0.013133	0.0000012	9.50E+07	0.019
Yemen	0.025998	0.0000028	4.14E+08	0.076

Appendix 1.C Normality Tests

Variables	Sample Correlation Coefficient ($r_{(Q)}$)
Area	0.849
Land use (% arable)	0.906
Population	0.486
Population growth rate	0.979
Life expectancy	0.965
Literacy rate	0.960
GDP	0.532
Televisions per capita	0.911
Airports per capita	0.779
Military expenditures	0.672
Military expenditures % GDP	0.828

Note: Critical Value = 0.9538
Significance Level $\alpha = 0.01$