

Predictability of Democratization: a Global Model

Northwestern University
College of Arts and Sciences
Department of Political Science
Statistics for Political Research
22 November 1999

Problem

Three waves of democratization have occurred in the modern world, each affecting a relatively small number of countries globally, and two thus far suffering from a reverse wave away from democracy. The first, with its roots in the American and French revolutions, existed between 1828 and 1926. The second started with World War II and existed between 1943 and 1962. The third wave began in Southern Europe in 1974 and following the fall of communism picked up speed.¹ It is as of yet unclear as to when this wave will end.

It is important for the political scientist to understand the reasons for the state's impetus for democratization, in order to make a meaningful prediction as to what states will democratize, and which that do will suffer from a backslide into an autocratic state. Prediction is particularly important for governments as politics and economics becomes increasingly globalized. Any state will be particularly concerned about a nation's level of democratization (defined as a continuous rather than a dichotomous variable) or hopes for democratization, before making any long term arrangements with it. An accurate prediction of the impending level of democratization for a state is crucial then to governments as we approach the twenty-first century.

No two of the processes towards democratization for the current 74 democracies were exactly alike. As a result, theories of transition vary from the over-simplistic that ensure their validity by their broadness to the over-specific that decline in usefulness due to their general non-applicability. For instance, Samuel Huntington, in his 1993 text *The Third Wave*, examines the movement towards democracy in each of the three waves independently, explains some non-universal commonalities amongst each wave, and particularly with the third wave is forced to divide the world into five prevailing patterns of democratization.² Global diversity has thwarted any universal formula for democratization and the author does not contend that he can overcome this.

The goal of this research then is to highlight specific traits amongst countries that, while not sufficient for democratization, are generally necessary to it. Such traits arguably include, but are not limited to, a strong economy, an educated populous, a low level of civil disorder, and a small military in respect to the size of the population.

¹ Huntington, Samuel P. *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press, Oklahoma: 1993. p.16.

² *ibid.*

Research Design and Hypothesis

The data used for the following research was compiled by Magoula Popi at North Carolina State University in a paper entitled, "Nations of the World Database."³ Data was gathered in 1980 and generally includes information for 111 states.

The Dependent Variable

Analyzing any variable's relation to democracy is complex because democracy in itself is unclear. Definitions of democracy vary between those that include strictly a procedural requirement, as that of Joseph Shumpeter, to those that also include substantive requirements, such as Huntington's definition. By using Popi's dataset, this research is dedicated to the index of democratization created in the dataset. Unfortunately, this is a relatively weak statistic. It only takes into consideration two empirical variables, both devoted to a procedural definition of democracy: the share of the smaller parties and independence of the votes cast in parliamentary and/or presidential elections and on the degree of electoral participation.⁴

This will have a devastating effect on the value of this research. The variable solely includes measures of contestation and participation but fails to include any measure of civil and political rights, responsiveness of political rulers, court bias and function, or media censorship. As a result, this research would not be accurate in claiming any correlation with democracy. In order to increase the accuracy (however decreasing the scope), the dependent variable will not be referred to by the misnomer, democratic index, but as a procedural democratic index (PDI).

Economic strength

It has been argued that for numerous reasons, a growing economy is correlated with a rise in democracy. Huntington argues that this occurs due to many reasons including:⁵

- 1) the rise of the middle class which holds the numbers and capability to call for a responsive government
- 2) the availability of greater resources for distribution among social groups and hence facilitates accommodation and compromise
- 3) an increase in the level of education in the society

³ Janda, Kenneth. *Statistical Analysis for Political Research: Polity – Nations of the World Database*. <http://janda.org/c10/data%20files/polity.html>. November 15, 1999.

⁴ *ibid.* p.5.

⁵ Huntington, Samuel P. *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press, Oklahoma: 1993. p.65.

This leads to the first hypothesis:

H1: The greater the per capita GNP, the higher the level of procedural democracy.

It is important though to distinguish between states that have abnormally high per capita GNPs due to natural resources such as Botswana with diamonds or Kuwait with oil, and states that acquire high per capita GNPs due to industrialization. The rise of the middle class, often the root of democratic leanings, is not associated with states such as Kuwait and Botswana. This leads to the second hypothesis:

H2: The greater industry's percentage of a state's GDP, the higher the level of procedural democracy.

Educational Level

A high level of education is correlated with a rise in democracy. Although it is true that education is strongly correlated with economic development as noted by Huntington, there are countries in which this is not the case. Countries that have a high level of economic development but lack industrial development generally have low levels of literacy. Additionally, there are some nations with higher levels of literacy but lack economic development due to numerous reasons including, but not limited to, war, unfavorable trade practices, and colonization. For these reasons, the third hypothesis is:

H3: The greater the male literacy level, the higher the level of procedural democracy.

Male literacy was used here as opposed to male and female literacy rates because females are often kept from educational resources as a social custom, and it is the author's contention that including females then would simply weaken the correlation.

Civil disorder

Countries with high levels of civil disorder are generally those in which the people do not have a peaceful outlet to express their discontent. This leads to the fourth hypothesis:

H4: The greater the level of civil disorder(CDI), the lower the level of procedural democracy.

Size of the military in respect to the populace

States that are autocratic generally have a large military presence at home to maintain order. Democratic states generally do not have this need. The dataset provides a soldier to civilian ratio (member of the armed forces per 1000 inhabitants). The fifth and final hypothesis is:

H5: The larger the soldier to civilian ratio, the lower the level of procedural democracy.

However, given that the dataset includes all armed forces in this ratio, rather than armed forces within the states borders, this hypothesis may prove incorrect.

Data Analysis

Histograms were run on each of the variables to determine the level of skewness. The procedural democratic index, the civil disobedience index, and the soldier to civilian ratio were each highly skewed. Base 10 logarithms were used in the analysis to normalize these distributions. Had these variables been analyzed with high skewness levels, the correlations may have been artificially high thus invalidating the results. Only 49 states were applicable for the study as a whole due to various missing cases.

TABLE 1: Correlations Matrix of Independent Variables and Procedural Democratic Index (FDI)

	FDI	Log of GNP per capita	Industry's % of GDP	Male Literacy Rate	Log of CDI	Log of Soldier to Civilian Ratio
FDI	1.00	0.66	0.29	0.66	0.07	0.07
Log of GNP/capita		1.00	0.48	0.78	0.05	0.46
Industry's % of GDP			1.00	0.44	0.23	0.07
Male Literacy Rate				1.00	0.05	0.35
Log of CDI					1.00	0.17
Log of Soldier to Civilian Ratio						1.00

The correlation matrix shows two significant findings. First, the logarithm of the GNP per capita, the industries percentage of GDP, and the male literacy rate all hold high product-moment correlations with the procedural democracy index (PDI). However, all of these independent variables have a high product-moment correlation with each other, therefore limiting the explanatory value of each variable and increasing the plausibility that either the logarithm of the civil disobedience index or the logarithm of the soldier to civilian ratio will obtain a greater explanatory value under multiple regression than the product-moment correlation suggests.

TABLE 2: REGRESSION EQUATION WITH DEMOCRATIC INDEX AS THE DEPENDENT VARIABLE

Variable	b Coefficient	s.e. of b	Beta Coefficient	t-value	Significance
Logarithm of GNP per capita	0.65	0.21	0.52	3.08	0.00
Male literacy rate	0.01	0.01	0.35	2.20	0.03
Logarithm of Soldier to civilian ratio	-0.58	0.23	-0.29	-2.54	0.15
(Constant)	-1.78	0.46	x	-3.86	0.00

Multiple R = 0.74 R Square = 0.55 Adjusted R Square = 0.55 Standard Error = 0.49 N = 49

The high correlation amongst the logarithm of the GNP per capita, the industry's percentage of GDP, and the male literacy rate noted in Table 1 has been factored out in the multiple regression. As a result, GNP per capita, male literacy rate, and the measure of civil disobedience are, together, the most significant in predicting a level of procedural democracy. The two insignificant variables, industry's percentage of GDP and the civil disobedience index, with t-significance levels of 0.30 and 0.47, respectively, have been excluded in the equation. Had they been included, the adjusted r^2 of 0.52 would have decreased.

The multiple regression yields an equation of the following:

$$PDI = -1.78 + 0.65(\text{Log of GNP/capita}) + 0.01(\text{Male Literacy Rate}) - 0.58 (\text{Log of Soldier to Civilian Ratio})$$

This implies that, for example, for every one unit increase of the logarithm of GNP per capita, there will be a 0.65 increase in the PDI, holding all other variables constant. The F-statistic is 18.32 with a significance of 0.00. This implies that the differences noted among states does not simply reflect random error in sampling and grants further validity to the results

Hypothesis 1 (the greater the GNP, the higher the level of procedural democracy) can be accepted. This variable has the highest Beta level, 0.52, and the lowest significance level confirming it to be the most influential independent variable.

Hypothesis 2 (the greater industry's percentage of a state's GDP, the higher the level of procedural democracy) must be rejected in multivariable regression. There is a relatively strong product-moment correlation between this variable and the PDI (0.29) but the majority of its was absorbed by the GNP per capita variable yielding a significance level of only 0.29.

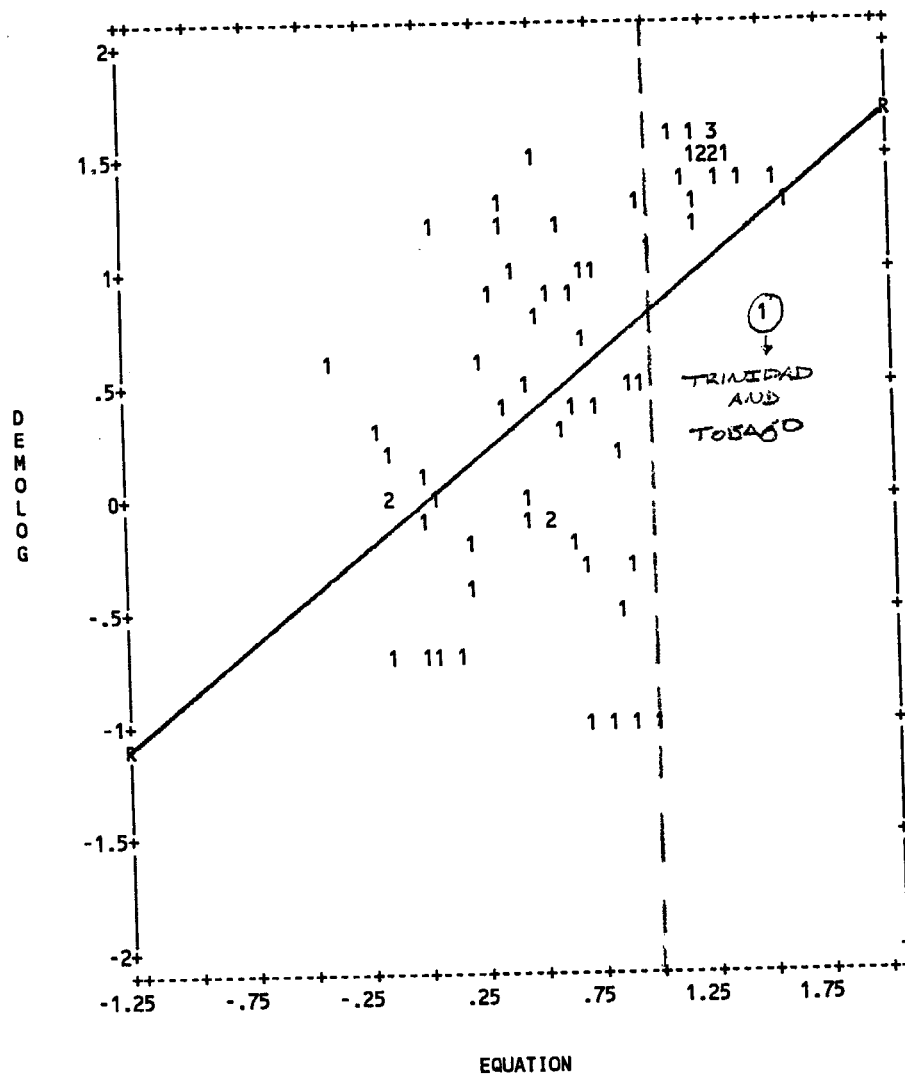
Hypothesis 3 (the greater the male literacy level, the higher the level of procedural democracy) can be accepted. Its effect was not fully absorbed by the GNP per capita variable.

Hypothesis 4 (the greater the level of civil disorder (CDI), the lower the level of procedural democracy) must be rejected due to a significance level of only 0.47.

Hypothesis 5 (the larger the soldier to civilian ratio, the lower the level of procedural democracy) can be accepted. After holding constant for the GNP per capita and the male literacy rate, this variable becomes significant (0.02) with a Beta level of -0.29 , implying that it has the least impact on the PDI of any of the three significant variables.

The equation stated above and the r^2 level are based upon global data. It is useful then to plot the predicted PDI against the observed logarithm of the PDI to see if the data is homoscedastic or heteroscedastic telling if the correlation is the same strength at any level of democratization.

FIGURE 1: Plot of Logarithm of PDI with Predicted PDI



69 cases plotted. Regression statistics of DEMOLOG on EQUATION:
 Correlation .53722 R Squared .28861 S.E. of Est .68480 Sig. .0000
 Intercept(S.E.) -.01015(.13815) Slope(S.E.) .85776(.16452)

The plot is heteroscedastic. The states that have a predicted PDI of greater than 1.00, with the exception of Trinidad and Tobago, have very low residuals. This implies that states that are first world democracies can be predicted well by the equation derived above. That is, states with a high GNP per capita, a high male literacy rate, and a low soldier to civilian ratio, are very likely to be procedurally

democracies. However the equation is not quite accurate in predicting the democratic level of a state that lacks any of such characteristics. States that are not strong democracies have very high residuals.

Summary and Conclusions

Political scientists and governments in particular are concerned with the level of democratization of a state at any given time and the likelihood that a state will increase on such scale. Multiple regression has shown that a high GNP per capita, a high male literacy rate, and a low soldier to civilian ratio, are significant in predicting that a state will be a democracy. Unfortunately, these variables yield a relatively poor prediction for democratizing countries, regardless of the high r^2 level.

The goal was to highlight specific traits that are necessary, but not sufficient for democracy. The research was a success in two ways. First, all states that are democratic were found to have such traits. Second, many states that have some, but not all, of these traits are not democracies. Therefore, it can be argued that a strong economy, an educated populace, and a small military to civilian ratio are necessary for a democracy, but not sufficient for it.

Procedural problems with this research include missing data, the inadequate democratic index produced by Popi, and data collected over an inadequate period of time. Further research would be well directed to divide democratized and democratizing countries in two groups and to run multiple regressions on them separately in order to provide more accurate statistics based on the different samples.

Paper Title: Predictability of democratization: A Global Model

The Problem (worth 3 points):

Score: 2

The problem is stated briefly while most paragraphs were devoted to the description of "democratic waves", rather than the possible causes of them.

Research design and hypothesis (worth 7 points):

Score: 6.5

Excellent job in dealing with the skewness of the independent variables, and you were right in paying attention to the intercorrelation between them.
The hypotheses are presented in a clear and systematic way.

Data analysis (worth 10 points):

Score: 8.5

Excellent job in explaining the heteroscedasticity of the plot, and your reasoning behind that makes a lot of sense as well.
Correct interpretation of multiple R squared, F, and B. But did you round up the adjusted R squared so that it equals to R squared?
The meaning of beta is not clearly explained for H1.

Summary and conclusion (worth 5 points):

Score: 5

A succinct summary of the research result.
The conclusion that the variables (that are significant) "yield a relatively poor prediction for democratizing countries" is particularly discerning and could have theoretical implication as well.

Total Score: 22