

To: President Elect Donald Trump
From: Isaac Haberman
Date: 12/14/2016
RE: Modeling the onset of Coups d'état

Executive Summary

Through my statistical analysis, I have built a model for predicting the onset of coups d'état. The State Department and the CIA will be able to practice better statecraft, better diplomacy, and better overall planning if they rely more heavily on statistical modeling and predictions. Using statistical analysis and insight garnered my review of academic papers, I tested a variety of logistic regression models and identified the best model. My chosen model has higher statistical accuracy, relies on fewer assumptions, and is easier to understand and recreate than the Ulfelder model. The State Department and The CIA should adopt my chosen model as their statistical model for predicting the onset of Coups d'état.

Importance of Coup Prediction

I have iterated below the importance of powerful statistical tools, including modeling and predictions, for the CIA and State Department:

- Better statecraft; with a greater understanding of events and actions that cause preferred results, the CIA and State Department can engage with other parties with greater certainty of how situations and conflicts will end. For example, the State Department can predict within a conflict which side will win and therefore which side the US should work with.
- Better diplomacy; with a powerful understanding of how parties will react to State Department actions, the State Department and CIA can commit to certain actions knowing full well what events will unfold. For example, the State Department could offer aid to a political leader, knowing whether or not they would take the aid and how it will affect their standing.
- Better planning; with more accurate predictions, the CIA and State Department can better predict how interventions and aid will affect situations. For instance, the CIA will know with more certainty whether arming select rebel groups in certain Middle Eastern countries allow for a preferred outcome or a faster end to the conflict.

Notes

- The original data set and model are from Jay Ulfelder, former research director of the Political Instability Task Force, a task force commissioned by the CIA.
- The word significant refers to statistical significance, a measure of the importance of a variable within a model.
- Probabilities range from 0 to 1, with 0 as an absolute non- occurrence and 1 an absolute occurrence.

- Predictions represent chances of coups. Simply said, the countries I have identified most likely to have a coup does not mean they will absolutely have a coup, rather they are most likely to have a coup.
- I will refer to the data set, including my additions, as the Ulfelder data set. I split the data set into two:
 - CW: full data set with observations from 1960 to 2014
 - Post: data set with observations after the end of the Cold War, 1991 to 2014
- For the purpose of this paper, I will refer to four models:
 - Ulfelder Model (CW): The original model published by Jay Ulfelder
 - Ulfelder Model (Post): The Ulfelder model run on post-Cold War data
 - Habermodel (CW): My chosen model for the full data set
 - Habermodel (Post): My chosen model for the post-Cold War data
- The Ulfelder data set extends only to 2014, therefore predictions for 2016 could not be generated.
- All models are predicting coups, the dependent variable, as defined by Jay Ulfelder in his original report¹.

Data

For my modeling, I used the aforementioned Ulfelder data set with the inclusion of the data set from *Determinants of the Attempting and Outcome of Coups d'état*. After reasoning through the data, I separated the data into two data sets, the full data set, CW, and a data set with only observations after the end of the Cold War, Post. I chose to use two data sets and create two models, as my Cold War indicator proved to be a significant predictor in all the models I tested. I was unable to identify what the significant difference between pre-and post-Cold War coups, so I made the two data sets and models to allow those interested to choose whichever they prefer.

Ulfelder Model

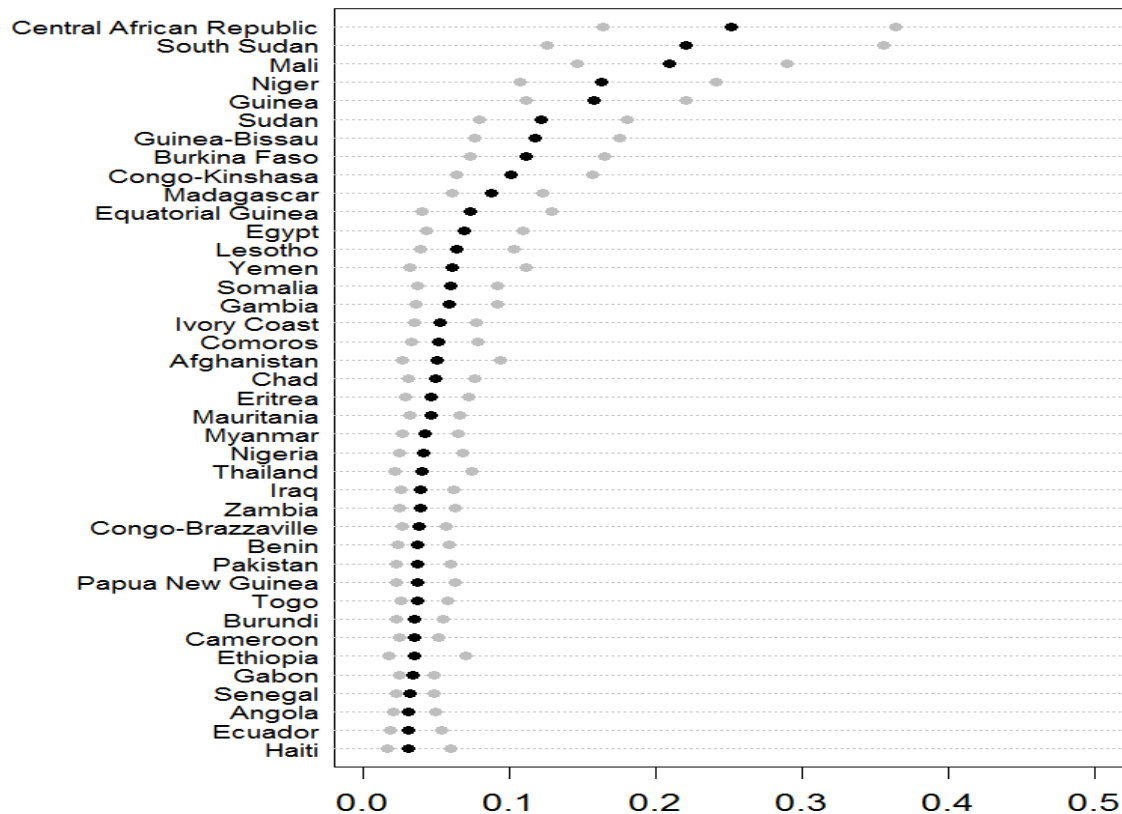
Below is the summary of the original Ulfelder model. The Ulfelder model is a logistic regression model run on the CW data set that produces an average accuracy of 78.82%. As can be seen below, Ulfelder has measures of government, poverty, country age, country makeup and world stability. The largest coefficients, excluding the intercept, were infant mortality rate and coup attempts in previous five years, both of which were predictors in my later models.

Variable	Coefficient	Std. Error	P-Value
Intercept	-3.08	0.68	5.4e-6
Colonized by Britain	-0.10	0.21	0.65
Colonized by France	0.28	0.21	0.19
Colonized by Spain	0.41	0.26	0.11
Logged age of country	-0.09	0.08	0.27
Infant mortality rate	0.74	0.15	1.3e-6

Coup attempts previous 5 years	0.94	0.17	3.3e-8
Coups globally	0.07	0.23	0.77
Coups Regionally	0.10	0.14	0.47
Previous annual GDP growth rate	0.41	0.16	0.01
Anocracy	0.01	0.23	0.95
Autocracy	0.07	0.24	0.77
Democracy	-0.41	0.36	0.26
Durability of regime	-0.21	0.08	0.01
Ethnic Elitism	0.19	0.16	0.23
Election year	0.18	0.17	0.28
Violent civil conflict	0.25	0.18	0.16
Cold War indicator	-0.71	0.22	0.00

Below are the Ulfelder coups predictions for 2015. The black dot represents the probability of a coup occurring, while the grey dots represent the 95% confidence interval. As can be seen below, none of Ulfelder's predictions are greater than 20%, indicating that there are very few coups predicted for 2015. Later in this memo, we will compare his predictions with the predictions of the Habermodel (Post).

Risk of Any Coup Attempts in 2015 (Top 40)



Variables to Test

I reviewed four academic papers on coups d'état modeling and summarized the significant predictors the authors had used, highlighting their respective theories:

- **Toward a Structural Understanding of Coup Risk²; Aaron Belkin and Evan Schofer**
Belkin and Schofer's article, models coup risk based on government structures, society, political culture, and state-society relations. Belkin and Schofer's final model has the following significant predictors:
 - Coup Risk measured by Bueno de Mesquita
 - Ln(GDP per capita)
 - South America, dummy
 - Central America, dummy
 - Regime type, binary

- **Revisiting Economic Shocks and Coups³; Nam Kyu Kim**
Kim's article analyzes the relationship between coup risk and economic shocks, both permanent and transitory. I have listed the significant variables from Kim's model below:
 - GDP growth previous year
 - Ln(GDP per capita) previous year
 - Interstate War previous year, binary
 - Coup previous year, binary
 - Rainfall deviation previous year x Agriculture
 - Temperature deviation previous year
 - Oil price shock previous year

- **Determinants of the Attempting and Outcome of Coups d'état⁴; Jonathan Powell**
Powell investigates coup-proofing of militaries to lower the chance of structural coups. The significant variables from Powell's model are listed below:
 - Expenditure per soldier
 - Military size in personnel
 - Paramilitary binary variable
 - Counterbalancing measured by Belkin and Schofer
 - Domestic Instability
 - Regime type measured by Polity score
 - Military regime binary variable

- **The Predictability of coups d'état: A model with African data⁵; Robert W. Jackman**
Jackman's paper explores the structural determinants of coups in Africa. Jackman's final model includes:
 - Social Mobilization
 - Size of largest ethnic group
 - Winning party percentage
 - Electoral turnout
 - Winning party percentage * turnout
 - Ethnic group size * winning party percentage
 - Ethnic group size * turnout

- Ethnic group size * winning party percentage * turnout

I was able to substitute for most of the significant variables from these papers in my modeling.

Model Selection

Using both data sets, I tested models until I reached a model that had both a higher average accuracy than the corresponding Ulfelder model and had predictors that fit within the academic papers I had read. Below I have listed all the variables I tested and the variables I used in my final model, the Habermode.

Variables	Ulfelder	Intermediate Models	Habermode
Colonized by Britain	*	*	
Colonized by France	*	*	
Colonized by Spain	*	*	
Post-Cold War (Included in CW data set)	*	*	*
Logged age of country	*	*	
Infant mortality rate	*	*	*
Coup attempts previous 5 years	*	*	*
Coups globally	*	*	
Coups Regionally	*	*	
Previous annual GDP growth rate	*	*	
Anocracy	*	*	
Autocracy	*	*	
Durability of regime	*	*	
Ethnic Elitism	*	*	
Election year	*	*	
Violent civil conflict	*	*	
Authoritarian Type: Party		*	*
Slow Growth		*	*
Central America Dummy		*	
South America Dummy		*	
Military size in personnel		*	
Expenditure per soldier		*	
Paramilitary binary variable		*	
Counterbalancing measured by Belkin and Schofer		*	
Authoritarian Type: Personal		*	
Authoritarian Type: Monarchy		*	
Authoritarian Type: Military		*	
Leaders Tenure		*	
Authoritarian Type: Party * Leaders Tenure		*	
Authoritarian Type: Party * Durability of Regime		*	
Authoritarian Type: Party * Slow Growth		*	
Authoritarian Type: Party * Country Age		*	

Model	Accuracy
Ulfelder (CW)	78.82%
Habermode (CW)	81.31%

Ulfleder (Post)	78.00%
Habermode (Post)	83.25%

To the left is a table of accuracies summarizing the average accuracy, over the entire data sets, of the four final models, the two Ulfelder models and the two Habermode models. The higher the accuracy, the better the model is at predicting coups.

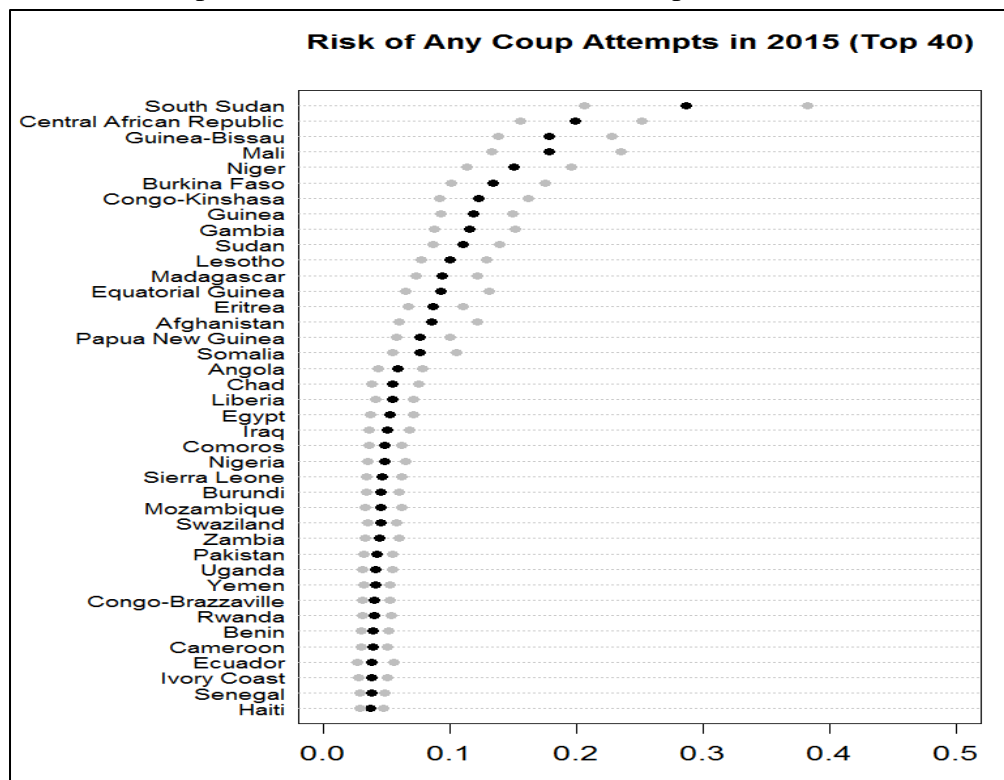
Chosen Model Summary

Below and to the right is the summary of the Habermode (Post). The Habermode (Post) is a logistic regression model run on the Post data set that produces an average accuracy of 83.25%. As can be seen below, the Habermode is a much simpler model than the Ulfelder models as well as some of my intermediate models. There are similarities between this model and the Ulfelder models; the two largest coefficients are infant mortality rate and coups attempts in the past five years. There is a measure of economic stability, and there is a measure of government type, echoing many of the predictors from the Ulfelder models.

Variable	Coefficient	Std. Error	P-value
Intercept	-3.82	0.21	<2e-16
Infant Mortality Rate	1.02	0.19	1.9e-10
Coup attempts in previous 5 years	1.00	0.25	8.4e-5
Authoritarian Party	-0.84	0.34	0.03
Slow Growth	-0.20	0.06	0.00

Below, are the Habermode coups predictions for 2015. Like the previous dot plot, the black dot represents the probability of a coup occurring, while the grey dots represent the 95% confidence interval. Again, like the Ulfelder predictions, most of the Habermode predictions are rather small, none greater than 30%.

Interestingly, while there is some variation between the predictions, most of the countries on the dot plot are in similar positions to those of Ulfelder predictions. However, the Habermode confidence intervals appears smaller the Ulfelder ones,



especially as the probabilities approach 0.

Recommendations

Based on my findings, I recommend the following:

The Habermodel, either the Post or the CW, should be used as the standard model for predicting the onset of coups for the relevant government agencies. I have shown above that the Habermodels are the most accurate and simpler than the Ulfelder models, and should, therefore, be used where appropriate.

Modeling and predicting begin to be used regarding other forms of governmental failure. Coups are a specific form of governmental failure; however, analysis should be done to model and predict other forms of governmental failure, such as revolutions or civil wars.

Government agencies begin standardized data collection, as the work of analysts can be simplified with standardized data sets. Almost all the data used for this work was created and managed by private citizens through academia.

Future Work

While I was able to develop a model that performed better than the Ulfelder Model (CW), the model can be improved. I believe further effort can be placed on developing better measures for predicting coups as well as developing a more complete data set that does not rely on incomplete, private data sets. For example, there may well be better measures of a country's economic stability and healthcare than the ones I chose. However, I used what was available to me in my data set. Additionally, I used Ulfelder's definition of coups; it's possible that a more stringent or a laxer definition yields different results.

Sources

¹ Ulfelder, Jay. "Coups Forecasts for 2014." *Dart-Throwing Chimp*. N.p., 2014. Web. 15 Dec. 2016.

² Belkin, A., and E. Schofer. "Toward a Structural Understanding of Coup Risk." *Journal of Conflict Resolution* 47.5 (2003): 594-620. Web.

³ Kim, N. K. "Revisiting Economic Shocks and Coups." *Journal of Conflict Resolution* 60.1 (2014): 3-31. Web.

⁴ Powell, J. "Determinants of the Attempting and Outcome of Coups D'etat." *Journal of Conflict Resolution* 56.6 (2012): 1017-040. Web.

⁵ Jackman, Robert W. "The Predictability of Coups D'état: A Model with African Data." *Am Polit Sci Rev American Political Science Review* 72.04 (1978): 1262-275. Web.