Methods and Initial Results

Sanittawan Tan

21 May 2019

1 Data

This section first discusses the challenges around measuring and classifying countries which are experiencing or have experienced democratic backsliding. It then explains the data collection process and the final data set that is used for analyses in this paper.

Although there is no scholarly consensus on how to measure or classify countries that are experiencing democratic backsliding yet, existing literature shows that there are two common approaches, both employing proxies or indices. The first approach is to use a reduction in Polity IV scores. The second approach is classifying countries based on the reduction in Varieties of Democracy (V-DEM) scores. In the second approach, Gottlieb et al. (2018) classify countries which are suspected to be experiencing democratic backsliding if a country's liberal democracy index as measured by V-DEM at time t is lower than that of time t-1.

As discussed in the previous section (literature review), there is still a major

gap between theoretical ground and the current measurement of democratic backsliding. Put differently, the current measurement of democratic backsliding does not necessarily comport with its theory. In fact, measurement is particularly challenging since democratic backsliding is more difficult to observe and its invisibility affects the coding and analysis of the country cases. Researchers may encounter measurement error which could jeopardize the credibility of the findings. Lueders and Lust (2018) found that there is little agreement among political scientists when conceptualizing and measuring regime changes. The measurement debate is by no means the scope of this paper; however, it is crucial to point to the current debate in the field.

As other scholars have employed Polity IV and V-DEM scores as proxies for the health of democracy, this paper aims to contribute to the debate by employing an alternative measure: Freedom House's country ratings as proxies for classifying regimes and selecting cases of democratic backsliding. A larger goal in play is to eventually test the discrepancies across sets of cases that were selected using different measurement which is a future project. This paper classifies a case as experiencing democratic backsliding if the Freedom House classifies the regime as "Partially Free." The Freedom House states in their methodology section that "...countries whose combined average ratings fall between 3.0 and 5.0 are Partly Free..." The overall score is obtained by combining political rights and civil liberties scores which range from 1 to 7. I argue that the Freedom House's rating may be a good proxy for indicating democratic erosion because in order to obtain the overall scores in the middle range, it is more likely that the country was rated in the middle range (3-5) for both political rights and civil liberties. It is less likely that a country that receives a rating of 6 or 7 in political rights (which means a country is ruled by authoritarian regime or

with severe government oppression will receive a score of 1 or 2 in civil liberties which means a country enjoys a wide range of civil liberties and freedom, hence receiving an overall score of 3.5 to 4). Lueders and Lust were partially correct when criticizing the Freedom House's aggregation (738). They state that there is a "multitude of logically possible combinations of individual scores [that] may yield the same rating on the Political Rights [or Civil Liberties] at the question level..." However, they may have overlooked the fact that each question is fairly correlated. Thus, the actual possible combinations for arriving at a certain level of ratings may be lower than what they expected.

When the Freedom House classifies a country case as partly free, it implies that such country case is not yet a full-blown authoritarian, but its democracy is highly flawed. This implication is consistent with most literature on democratic backsliding including Waldner and Lust (2018) and Bermeo (2016). In addition, Bermeo (2016) and Varol (2015) observe that in some cases erosion (or stealth authoritarianism in Varol's conception) was initiated by the executive branch. Regardless of how erosion occurs, Lust and Waldner and Bermeo tend to agree that the result is weakened political institutions which are supposed to serve as checks and balances against the government. Thus, this paper's hypothesis is that deterioration in the quality of political institutions should correlate with or be a good predictor of democratic erosion.

In order to test the hypothesis about democratic backsliding, I constructed a data set that combines three following data sources.

- Freedom House country and territory ratings and statuses (1973-2018)

- Boix-Miller-Rosato Dichotomous Coding of Democracy (1800-2015)
- V-Dem Country-Year Dataset v9 (Full 450 V-Dem indicators and 59 other indicatiors)

The two dependent variables of interest are (i) whether or not the country is experiencing or has experienced democratic backsliding and (ii) Freedom House's country status classification. Both are retrieved from the Freedom House country and territory ratings and statuses (1973-2018).

Regarding independent variables, I carefully selected a wide range of indicators from Boix-Miller-Rosato Dichotomous Coding of Democracy (1800-2015) (hereafter "BMR") and V-Dem Country-Year Dataset v9 (Full 450 V-Dem indicators and 59 other indicatiors) (hereafter "V-DEM"). Variables "democracy breakdowns" which records a country's number of democratic breakdowns and "democracy duration" which records the number of consecutive years of current regime type are selected from BMR. I selected 65 V-DEM indicators from 12 categories including those that entail the quality of the executive, the legislature, and the judiciary. Please refer to the attached Appendix 1 for a full list of selected indicators.

In terms of organization of the data set, I follow the tradition of scholarly work in regime change literature by coding up longitudinal country-year data set. In other words, each country and year is treated as independent case or observation in the data set. The final data set which is used for the analysis in the subsequent section has 79 columns and 1,946 observations. The number of columns increases from 67 because some of the independent variables are categorical and have to be one-hot encoded. The breakdown of country-year cases are shown in table 1 and 2. (The full

data set can be found in the data directory with the name "democracy.csv".)

Table 1: Freedom House's classification breakdown

Category	Number of Cases
Not Free (0)	478
Partly Free (1)	588
Free (2)	880

Table 2: Cases of Democratic Erosion

Category	Number of Cases
No Democratic Erosion (0)	1358
Democratic Erosion (1)	588

2 Model

In order to test the hypothesis, this paper employs Random Forest model to predict if a country case is experiencing or has experienced democratic backsliding. Because Random Forest model does not lend itself for an easy interpretation, we also leverage feature importance which is a model-agnostic technique used for interpreting complex statistical learning models. If the hypothesis were to be correct, feature importance method should show that indicators related to weakening quality of political institutions are important features of cases where democratic backsliding is present.

Since the data set is longitudinal, I model that the classification by the

Freedom House and whether or not democratic backsliding is present are results of deterioration in relevant indicators in the previous years. Hence, all independent variables are lagged by one year when used for prediction.

I primarily ran two models. The first model tries to capture the relationship between democratic backsliding and the selected indicators. The second model seeks to explain the important predictors of the Freedom House's classification into free, partly free, and not free statuses. The two following figures depict samples of decision trees built in each model. The next section discusses the preliminary results.

Figure 1: Sample tree from democratic erosion classification

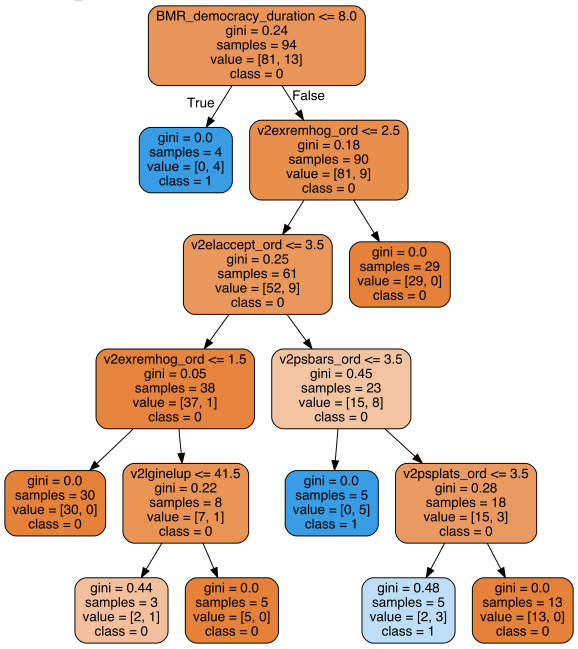
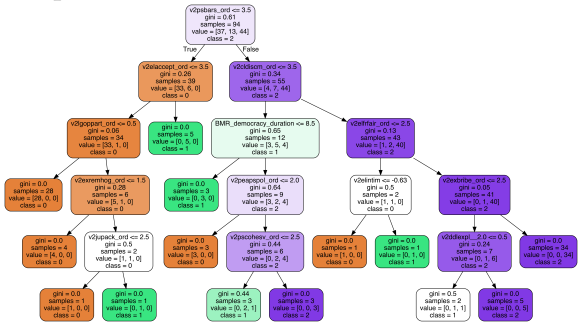


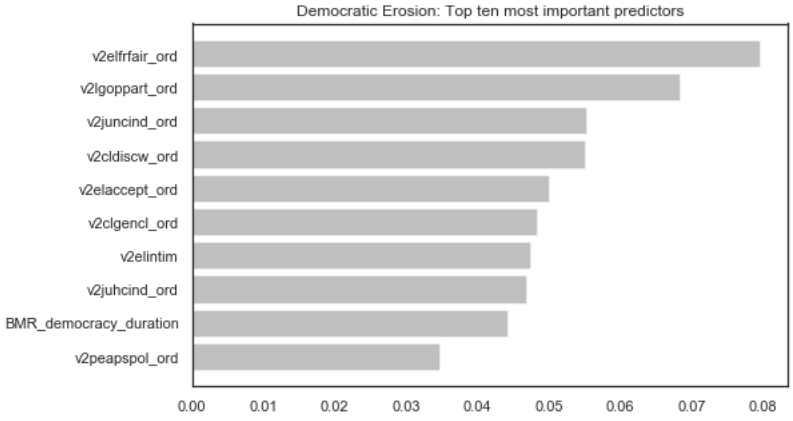
Figure 2: Sample tree from Freedom House status classification



3 Preliminary Results

According to the first model where we predict whether democratic erosion is present in a country case, the top ten features are as follows:

Figure 3: Top ten features - Democratic Erosion

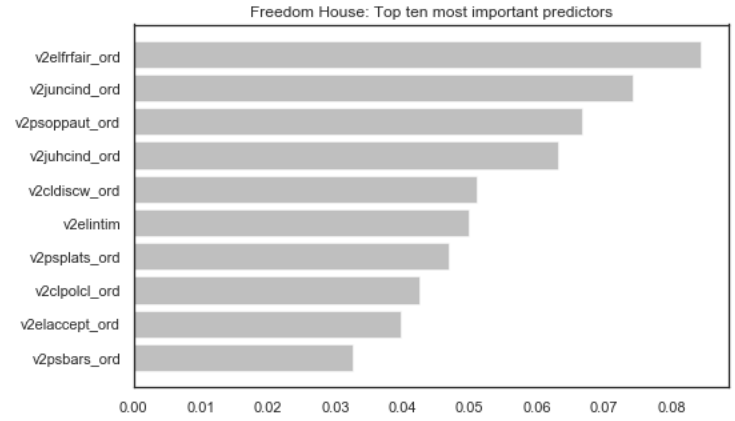


- Election free and fair
- Legislature opposition parties
- Lower court independence
- Freedom of discussion for women

- Election losers accept results
- Gender equality in respect for civil liberties
- Election government intimidation
- High court independence
- Democracy duration
- Access to public services distributed by political group

According to the second model where we predict the classification of Freedom House's country statuses, the top ten features are as follows:

Figure 4: Top ten features - Freedom House classification



- Election free and fair
- Lower court independence
- Opposition parties autonomy
- High court independence

- Freedom of discussion for women
- Election government intimidation
- Distinct party platforms
- Political group equality in respect for civil liberties
- Election losers accept results
- Barriers to parties

It seems that our hypothesis is not entirely correct based on the data. However, the features from the two models tend to be correlated. This is not surprising since our measure of democratic backsliding is derived from the Freedom House's ratings. The final section discusses the preliminary results and potential next steps for the paper.

4 Discussion and the Next Steps

The results from the model suggest that the quality of election, political parties and court independence seem to be good predictors of whether or not democratic erosion is present. The result does not seem to be consistent with our hypothesis. However, there are several potential problems to our methodology.

Firstly, Random Forest classifier does not handle missing values which causes us to drop observations that have at least one missing values from the data set. This

significantly reduces the number of observations from 1,946 to just 94 observations. The reduction in the number of observations certainly contributes to increased variance of the findings. To handle the problem of missing values, I plan to investigate the best way to impute missing data using techniques covered in class. Another solution is to reduce the number of predictors and focus on indicators that pertain to political institutions and elections. If there are missing values, imputation can be carries out.

Secondly, feature importance method involves certain level of randomness; as a result, the identified important features have high variance. This problem stems from the fact that data has to be shuffled prior to feature importance calculation. Different ways of shuffling could change the identified features. One solution to this is perhaps to repeat feature importance method many times and look for common key features that are identified in each iteration.

Thirdly, Random Forest classifier treats country-year cases as independent of each other. However, in reality, one could make a case that events that occur in the past affect the current situation. Hence, we should probably find a better model that grow trees chronologically. There seems to be some literature on the matter from bioinformatics field which deals largely with health data. Reviewing the methodology used in such field may help improve our predictability and get closer to the reality. I may not be able to carry this out within this quarter, but it is an area worth investigating.

Finally, the result may suffer from measurement bias since there is not yet an absolute or fool-proof measure of democratic erosion. The finding in this paper calls into question the existing measure and speaks to Lueders and Lust' concern that

regime change research findings are susceptible to the change in measurement.