

## Introducing Democratic Erosion

Democratic erosion, interchangeably democratic backsliding, is an emerging concept in regime change literature in comparative politics. Nancy Bermeo (2016) defines it as “the state-led debilitation or elimination of any of the political institutions that sustain an existing democracy.”

Democratic erosion differs from an abrupt suspension or breakdown of democracy via conventional means such as coup d'état because it underscores gradual and incremental changes usually undertaken by the incumbent government or the executive branch through perfectly legal means such as election manipulation, expansion of the executive power and suppressing press freedom.

Since the process is often subtle, detection and measurement pose a major challenge to scholars.

## Research Question and Hypothesis

Although there is no consensus on measurement and classification of cases where democratic erosion occurs yet, we would still like to learn something about the **predictors** of such phenomenon.

Many scholars observe that democratic erosion is associated with:

- **deteriorating quality of democracy**
- **weakened political institutions**

Thus, we hypothesize that, if this is the case, indicators related to the two factors could be predictors or important factors of democratic erosion. However, we do not attempt to study causality here.

## Data & Methods

We construct a longitudinal data set from 3 data sources. The data set consists of 179 countries from 2000 to 2017, totaling 3,478 country-year cases.

- Dependent variable is a dichotomous variable indicating if a particular country-year experienced democratic erosion. This is derived from the Freedom House's Freedom in the World country statuses.
- Independent variables are 60+ indicators of democratic quality and institutions collected from Varieties of Democracy (V-DEM) and Boix-Miller-Rosato (2018) lagged by 3 years.

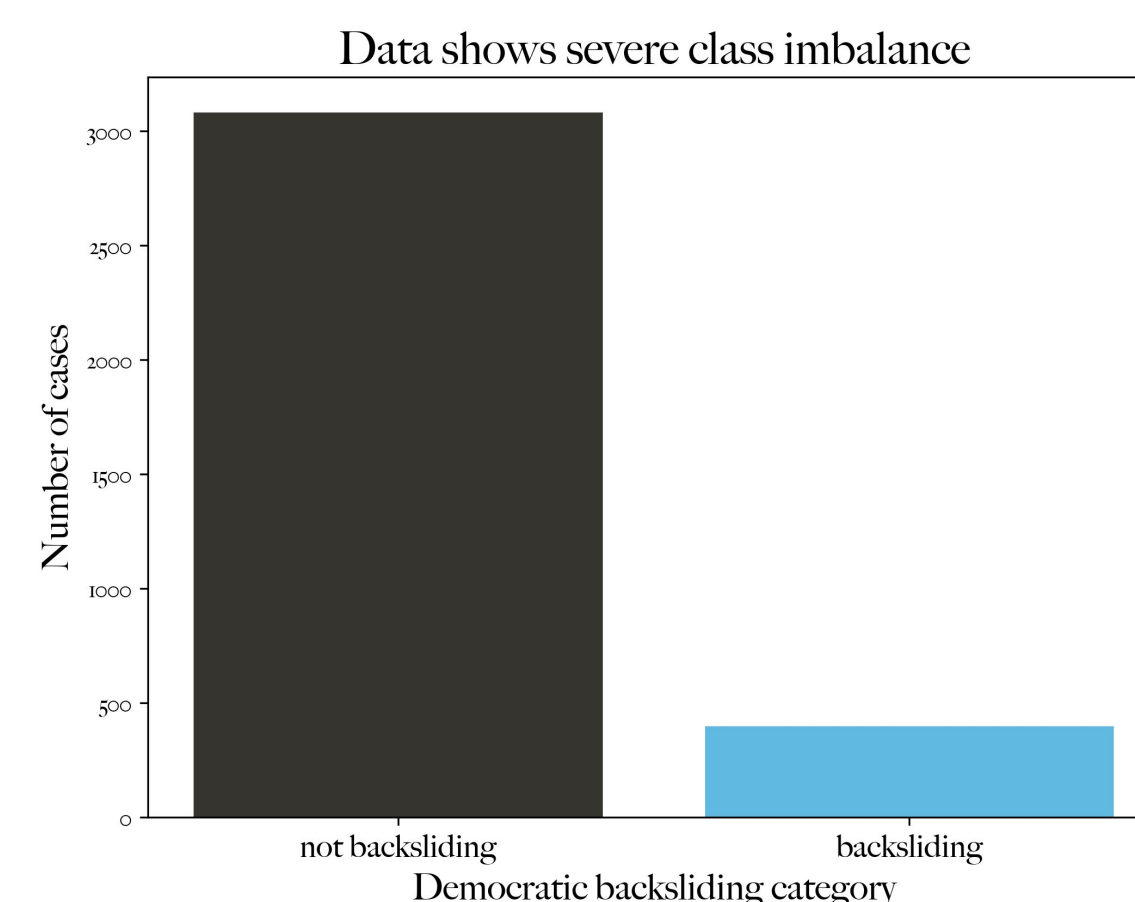
Our research goal is to use these indicators to predict cases of democratic erosion and non-erosion.

## Challenges of the Data Set

The data set that we compile poses three problems.

1. **Class imbalance** — As figure 1 illustrates, democratic erosion is much rarer than cases of non-erosion. To deal with this, we use **stratified sampling** which ensures the same class proportion in the training and test sets.

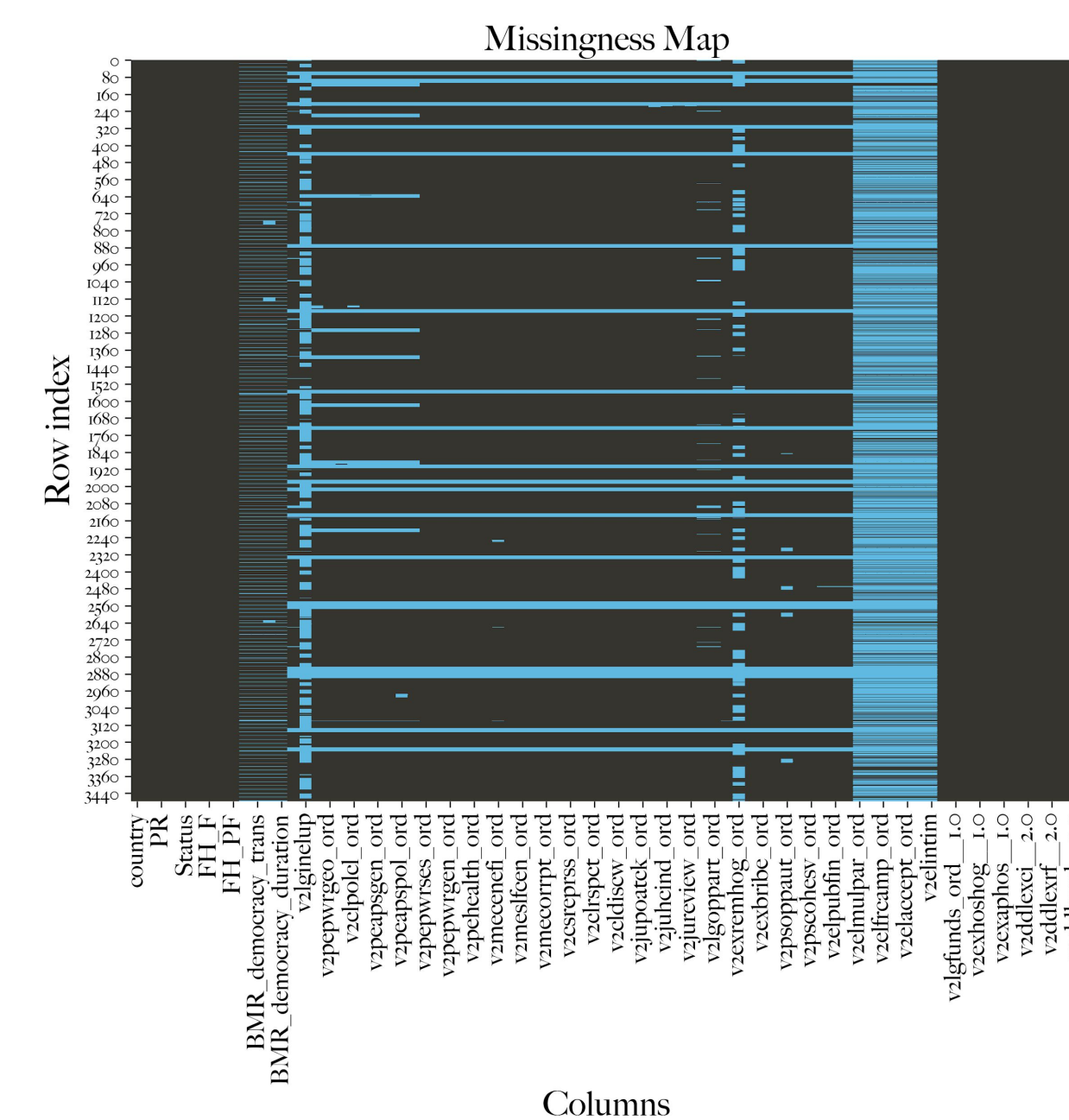
Figure 1: Imbalanced class



## Challenges of the Data Set (cont.)

2. **Missingness** — As figure 2 illustrates, many indicators are missing *not at random*. Thus, we use **XGBoost**, a boosting method that treats missing data as its own category, for prediction. However, columns with a lot of missing data are dropped regardless.

Figure 2: Mapping missing data



3. **Temporal dependencies** — Because the dataset is longitudinal, treating each observation independently is unrealistic in cross validation (CV). We use **rolling validation sets** instead of traditional CV approach.

## Leveraging XGBoost

Because the dependent variable is dichotomous, it can be thought of two different classes. Now the problem becomes classification. XGBoost was selected for the task because it deals with missing data and is well-known for its predictive power.

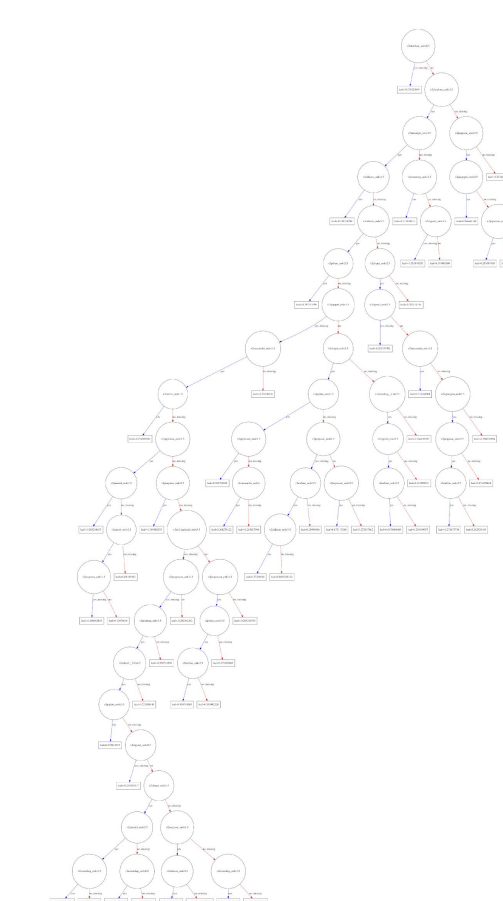


Figure 3: Sample tree

## Results

- The model achieves ~0.8 test set accuracy score
- Our real interest is interpreting the model and find out what the key predictors are. For this task, we use **feature importance** approach.
- However, feature importance involves permuting the dataset and it suffers from high variance. Thus, we implement 100 iterations.
- Important features appear to be indicators on governance, election and participation. The most interesting one is duration of democracy.

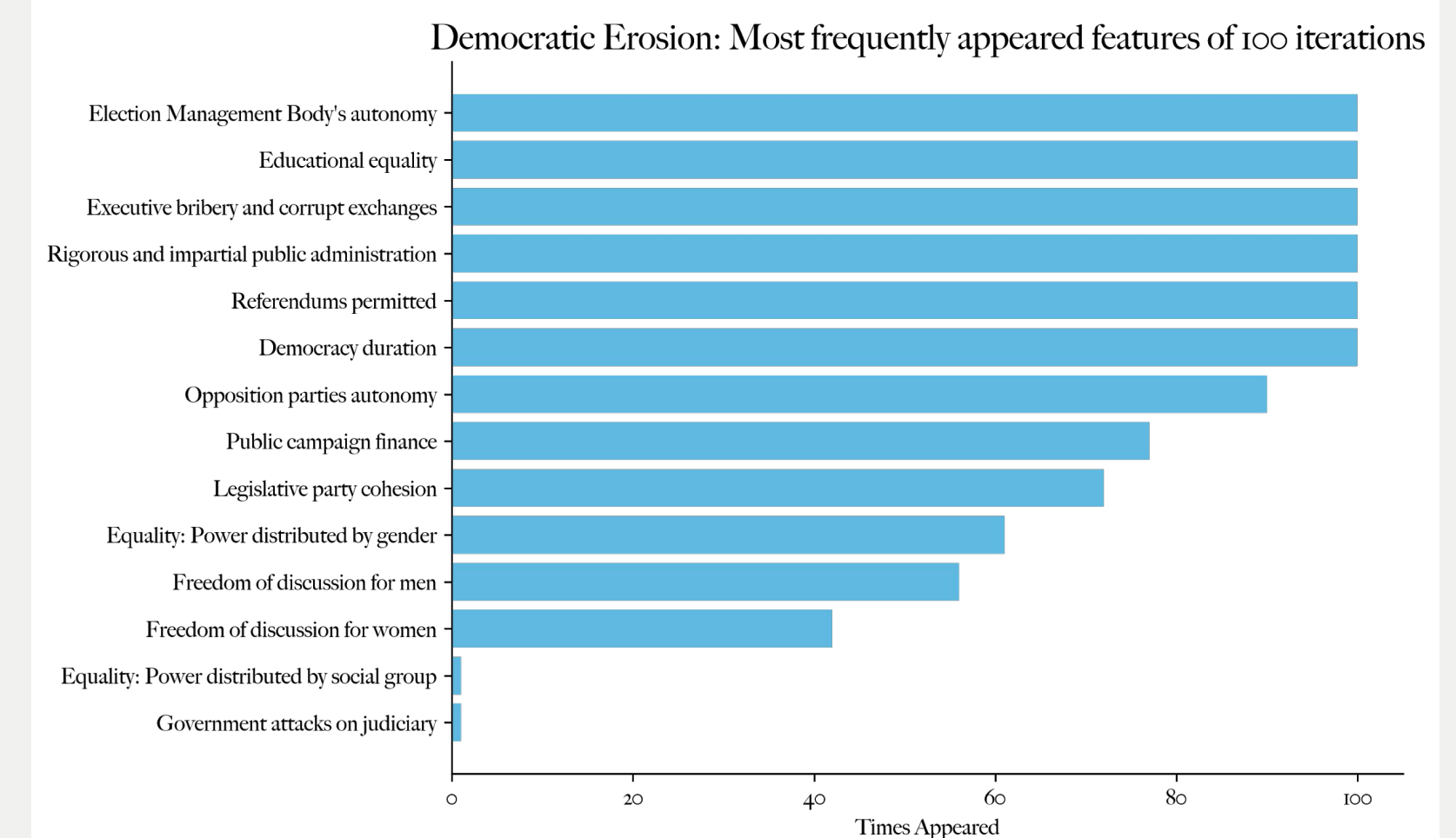


Figure 4: Importance features

## Key Takeaway # 1

It is surprising that indicators which reflect weakened institutions such as indicators on the court and the legislative branch do not appear in the top ten most important features. This could be because of imperfect measurement and modeling approach.

## Key Takeaway # 2

We need to develop a better measure with strong theoretical ground for categorizing cases that fall inside the realm of democratic erosion.