Varieties of Forecasts: Predicting Adverse Regime Transitions

Richard K. Morgan, Andreas Beger,
Adam Glynn

May 2019

Working Paper
SERIES 2019:89
THE VARIETIES OF DEMOCRACY INSTITUTE

UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

Varieties of Democracy (V-Dem) is a new approach to conceptualization and measurement of democracy. The headquarters—the V-Dem Institute—is based at the University of Gothenburg with 17 staff. The project includes a worldwide team with six Principal Investigators, 14 Project Managers, 30 Regional Managers, 170 Country Coordinators, Research Assistants, and 3,000 Country Experts. The V-Dem project is one of the largest ever social science research-oriented data collection programs.

Please address comments and/or queries for information to:

V-Dem Institute

Department of Political Science

University of Gothenburg

Sprängkullsgatan 19, PO Box 711

SE 40530 Gothenburg

Sweden

E-mail: contact@v-dem.net

V-Dem Working Papers are available in electronic format at www.v-dem.net.

Varieties of Forecasts: Predicting Adverse Regime Transitions*

Richard K. Morgan
Varieties of Democracy Institute
University of Gothenburg

Andreas Beger
Predictive Heuristics

Adam Glynn
Department of Political Science
Emory University

*This research project was supported by European Research Council, Grant 724191, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden; the Knut and Alice Wallenberg Foundation (PI: Staffan I. Lindberg) and the University of Gothenburg (E 2013/43), as well as internal grants from the Vice-Chancellors office, the Dean of the College of Social Sciences, and the Department of Political Science at University of Gothenburg.

Abstract

This article introduces the V-Forecast project, the forecasting intuitive of the Varieties of Democracy (V-Dem) institute. In this the initial year of the V-Forecast project, we provide two-year ahead forecasts of the risk of adverse regime transitions (ARTs) for 169 countries. ARTs are substantial movements of a country's regime towards more authoritarian governance, whether authoritarian reversals in a democracy, or further autocratization in an already nondemocratic country. Examples include Hungary and Poland over the past few years, which are prominent cases in a more widespread and worrying global trend over that effects a significant fraction of the world's population. Yet so far, there has been no public forecasting system for anticipating new ARTs and identifying countries most at risk. We describe an effort that forecasts ARTs – operationalized using the Regimes of the World (RoW) categorization – with an ensemble model that leverages V-Dem and several additional external data sources. Despite being rare events with a roughly four percent baseline chance over any two-year period, in test forecasts the model is able to achieve good accuracy.

1 Introduction & Executive Summary

In recent years, political elites in a number of third wave democracies – most notably, Hungary, Poland, the Philippines, and Turkey – have been systematically undermining important democratic norms and institutions (Bermeo 2016, Diamond 2015, Kurlantzick 2013, Levitsky & Ziblatt 2018, Lührmann & Lindberg 2019, Lührmann, Mechkova, et al. 2018, Mechkova et al. 2017). In fact, no less than 24 countries experienced some form of democratic erosion between 2007 and 2017, affecting one third of the world's population and "represent[ing] a massive reduction in the global protection of rights and freedoms" (Lührmann, Dahlum, et al. 2018, p. 6). While the effects are often more diffuse and less intense than those associated with other political phenomena such as civil conflict, democratic erosion arguably has a greater negative impact on more people worldwide and over a longer term. Therefore, developing models that can help identify countries at risk of democratic erosion is of tremendous importance. This article introduces the V-Forecast project, the Varieties of Democracy (V-Dem) institute's efforts at developing these models.

There are a number of forecasting efforts currently underway throughout the international relations and comparative politics research communities. These projects range from predicting the onset of civil and international conflict (Brandt et al. 2011, Hegre et al. 2019, 2013) and mass killings and atrocities (Goldsmith & Butcher 2018, Goldsmith et al. 2013, Woocher et al. 2018) to whether a country will experience an irregular leadership change (Beger et al. 2016, Ward & Beger 2017) and political instability, in general (Goldstone et al. 2010). For the most part, these forecasting efforts focus on estimating the potential risk that a country will experience some form of political violence. To our knowledge, the field lacks a comprehensive forecasting system looking specifically at democratic erosion. With the goal of developing a suite of forecasting models focused on phenomena related to democratic erosion, the V-Forecast project aims to fill this gap.

As an initial step, the V-Forecast project is focusing on estimating a country's risk of experiencing an *adverse regime transition* (ART) within a *two-year* window. We conceptualize ARTs as a decline in the democratic qualities of a country's political regime. These declines can coincide with violent events such as coups and internal conflict. The military coup in Thailand in 2014 and the civil conflict (and subsequent coup) in Mali in 2012 are clear manifestations of ARTs that occur through these more violent processes. An ART can also stem from an incumbent regime's repressive response to political protests, as was the case in Bangladesh in 2012 when the government used violence to suppress protests. Further, ARTs also capture the gradual erosion of democratic norms and institutions. The events that have unfolded in Hungary over the past few years – Prime Minister Orbán's attacks on judicial constraints on executive power and his curtailment of media freedoms – is an example of this type of ART.

We capture ARTs using the Regimes of the World (RoW) index, which classifies political regimes as either a closed autocracy, electoral autocracy, electoral democracy, or liberal democracy (Lührmann, Tannenberg, & Lindberg 2018). While we plan to explore additional operationalization strategies in subsequent iterations of this project, we currently operationalize *adverse regime transition* as a year-to-year decrease in the RoW index. That is, an ART occurs when a country moves down the RoW index from one year to the next. We forecast the risk that such an event will occur within a two-year window. We describe our outcome variable in more detail in Section 2.2.

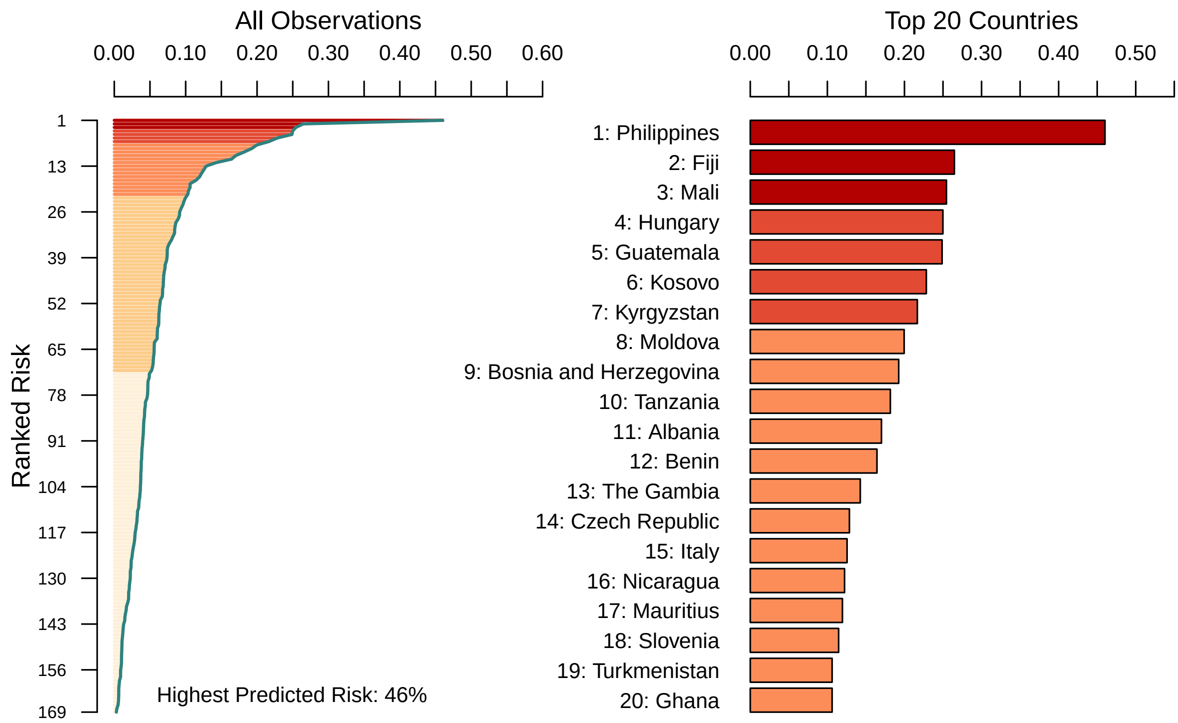
To produce our estimated risk forecasts, we use V-Dem data version 9 (Coppedge et al. 2019b, Lindberg et al. 2014) along with UN GDP and population data, ethnic power relations data (Vogt et al. 2015), coup event data (Powell & Thyne 2011), and armed conflict data (Gleditsch et al. 2002, Pettersson & Eck 2018), over 400 variables altogether. We use these data in three machine learning models: logit with elastic-net regularization, random forest, and gradient boosted forest. To help account for differences across these models, we use an unweighted model average ensemble. This is our preferred approach, as it helps smooth out our predictions while improving accuracy. Also, by using off-the-shelf machine learning models and an unweighted average of these models, this approach is also relatively simple and quite transparent. We detail our data set and outline our estimation techniques in Section 2.

In Section 3, we discuss our framework for evaluating model performance. In particular, Section 3 describes our 2×7-fold cross-validation procedure and our performance metrics, e.g., Area Under the Curve-Precision/Recall (AUC-PR). While aggregate measures of model performance are standard practice, they do not allow researchers to assess how randomness or other fluctuations on the data might be affecting model performance from year to year. In short, the accuracy of our yearly forecasts, which is dependent on the actual observed ARTs in a given year, are likely to deviate.

To assess how robust our models are to these fluctuations, we conduct a series of yearly test forecasts for 2011 through 2017. These test forecasts mimic the process we use to generate our "live" two-year-out forecasts, allowing us to infer the general performance of our 2019-2020 forecasts. Further, to evaluate whether groups of countries with similar predicted probabilities will experience ARTs at the rate implied by the group's risk scores and to assess whether our models are over- or under-predicting ARTs, we conduct a series of calibration tests and simulation experiments.

Based on industry standards for models and prediction problems of this nature (i.e., machine learning models for rare-events), our models preform remarkably well. Our ensemble model reports an AUC-PR score of 0.46 in our 2×7 repeated cross-validation procedure and an AUC-PR score of 0.39 in our set of yearly test forecasts. As a general benchmark of performance, an AUC-PR score that is higher than the observed frequency of events in the data is a signal that the model is an improvement over chance. With an

Figure 1: Estimated Risk of ART for 2019-2020 (Ensemble Model)



observed frequency of ARTs at roughly 4 percent, our ensemble model exceeds performance expectations.

Using the estimated probabilities for the top-ten cases for 2017-2018, we can use simulation techniques to determine how many ARTs we would have *expected* to occur within the set of top-ten estimates. Based on these simulations, we find that there was a 61 percent chance that at least three observations would have experienced an ART in our set from 2017-2018, with a 14 percent chance of five or more ARTs. The actual number of ARTs in the 2017-2018 top-ten estimated risk cases was five; thus, our ensemble model produced rather conservative estimates.

Looking now at our estimates for 2019-2020, the panel on the left in Figure 1 shows the range and spread of our risk estimates for 2019-2020, while the panel on the right focuses on the top-20 at-risk cases. Using these estimated probabilities, our simulation procedure suggests that there is a 75 percent chance that there will be at least two ARTs within the top-ten cases for this prediction window. Given the general performance of our ensemble model and these simulation results, a conservative estimate is that at least one of the top-ten should experience an adverse regime transition in 2019 or 2020. Section 4 provides a detailed discussion of our 2019-2020 forecasts.[1]

Nevertheless, it is important to note here that these forecasts are probabilistic. A high estimated risk does not mean that an ART will occur with certainty; similarly, a low estimate does not mean that an ART will not occur. Simply put, these are probabilities not certainties.

[1] A complete list of all 169 risk forecasts can be found in Table 7 in Appendix A

Moving forward, we plan to explore other ways of operationalizing ARTs. On current approach, which relies on the RoW classification scheme, is potentially risky. The uses of threshold cutoffs across a number of different V-Dem variables when determining a country's RoW class means we run the risk of capturing small changes in border cases rather than substantial and significant decreases in the democratic qualities of a country. We address this issue in our discussion of our outcome variable, Section 2.2, and in our concluding remarks.

In the near-term, we also plan to expand the scope of this project by focusing on distinct forms of democratic erosion as well as electoral violence. In particular, we are developing a series of thematic forecasting models that estimate the risk that there will be a decline in six dimensions of democracy: (1) elections, (2) civil society, (3) freedom of expression and media, (4) public corruption, (5) democratic inclusion, and (6) legislative and judicial constraints on the use of executive power. By developing forecasting models for a variety of autocratization phenomena and by making these predictions public, the V-Forecast project hopes to provide useful tools for policy-makers and aid agencies.

2 Methodology

We use country-years as our unit of analysis and limit our temporal frame to 1970-2018. We reconcile the differences between the V-Dem country-year set and the Gleditsch and Ward (GW) country-year set to facilitate the use of external data (e.g., coup data, UN population data, etc.).[2] This leaves 169 countries for our 2019-2020 forecasts. Our training and validation country-year set captures 7,754 observations.

2.1 Data

V-Dem version 9 is our primary data source; 417 of the 451 variables we have in our data set come from or are derivatives of V-Dem data. UN GDP and population data, ethnic power relations data (Vogt et al. 2015), coup event data (Powell & Thyne 2011), and UCDP's internal armed conflict data (Gleditsch et al. 2002, Pettersson & Eck 2018) make up the remaining 34 variables in our data set. We lag all variables one year and derive the first differences for a number of variables. In brief, we use data from 2011 to estimate the risk that an ART will occur in 2012 or 2013, for example. A complete list of variables as well as basic descriptive statistics can be found in Appendix E.

[2]We drop the following eight countries from the V-Dem set: São Tomé and Principe, Seychelles, Vanuatu, Palestine/West Bank, Palestine/Gaza, Somaliland, Hong Kong, and Zanzibar, as they lack coverage in the GW country-year set. Along with micro-states, we drop the following four countries from the GW set: Bahamas, Belize, Malta, and Brunei, as they lack coverage across the V-Dem county-year set. We also drop Bahrain from the GW set. While the V-Dem data provides some coverage of Bahrain, there are missing values for the entire series of Bahrain across a number of key indices; therefore, we excluded it. The number of countries in our data per year ranges from 137 to 169.

There are several missing values throughout the V-Dem data. For the most part, this missingness is concentrated around variables related to elections and the legislative branch. Most missingness is the product of four V-Dem coding rules: (1) when elections are permitted, election specific variables are not scored for the years between elections, (2) when elections are not permitted or when there is an interruption, all elections variables during that period are not recorded, (3) in the event of an interruption, key legislative-related variables are also not recorded,[3] and (4) legislative-related variables are not coded for the first year that institution was enacted.

When elections are permitted, we fill in missing variables by carrying forward the last non-missing value. If an election-related variable is missing because elections were not permitted or because there was an interruption, we set these observations to zero and create a dummy variable indicating whether the elections were, in theory, permitted. Similarly, when legislative-related variables are missing due to an interruption, we set these observations to zero and create a dummy variable signaling this change.[4] Finally, we back-fill all legislative variables that are missing because it is the first year of that there was a viable legislative branch. We maintain annotated code for all changes made to the original V-Dem v9 data.[5]

2.2 Adverse Regime Transitions

We operationalize *adverse regime transitions* using the Regimes of the World (RoW) index. This index classifies political regimes as either a closed autocracy, electoral autocracy, electoral democracy, or liberal democracy. To produce these classifications, the RoW index takes into account the quality of a country's electoral institutions (e.g., the presence of multiparty elections and the quality of these elections), its liberal characteristics (e.g., legislative and judicial constraints on the use of executive power), as well as the regime's record across various civil liberties indices (e.g., adherence to the rule of law and secure and effective access to the judicial system).[6] An adverse regime transition occurs when a country moves down this scale (going from an electoral autocracy to a

[3]According to the V-Dem Codebook, an "interruption" is typically the result of a coup, declared state of emergency, or military defeat (For election variables, see Coppedge et al. (2019a, p. 275), and for legislative variables, see Coppedge et al. (2019a, p. 46 and 132-35).

[4]In the Data Appendix E below these variables are named `is_elec` and `is_leg`.

[5]Our data management `R` code is available upon request.

[6]Specifically, the RoW index classifies countries according to the following criteria: "Electoral democracies score above 2 on the indicators for multi-party (`v2elmulpar_osp`) and free and fair elections (`v2elfrfair_osp`), as well as above 0.5 on the Electoral Democracy Index (`v2x_polyarchy`). Liberal democracy meets the criteria for Electoral democracy but also satisfy the liberal dimensions by a score above 0.8 on the V-Dem Liberal Component index (`v2x_liberal`), as well as a score above 3 on transparent law enforcement (`v2cltrnslw_osp`), access to justice for men (`v2clacjstm_osp`) and women (`v2clacjstw_osp`). Electoral autocracies fail to meet one or more of the above-mentioned criteria of electoral democracies, but subject the chief executive and the legislature to de-jure multiparty elections as indicated by a score above 1 on the V-Dem multiparty elections indicator (`v2elmulpar_osp_leg/_ex`). Closed autocracies do not satisfy the latter criterion" (Coppedge et al. 2019a, p. 219).

Table 1: Year-to-Year Transition Frequency Table – 1970-2018

		To:			
		Liberal Democracy	Electoral Democracy	Electoral Autocracy	Closed Autocracy
From:	Liberal Democracy	20.45%	**0.30%**	**0.00%**	**0.00%**
	Electoral Democracy	0.52%	19.18%	**1.01%**	**0.12%**
	Electoral Autocracy	0.03%	1.70%	25.05%	**1.02%**
	Closed Autocracy	0.00%	0.09%	1.92%	28.63%

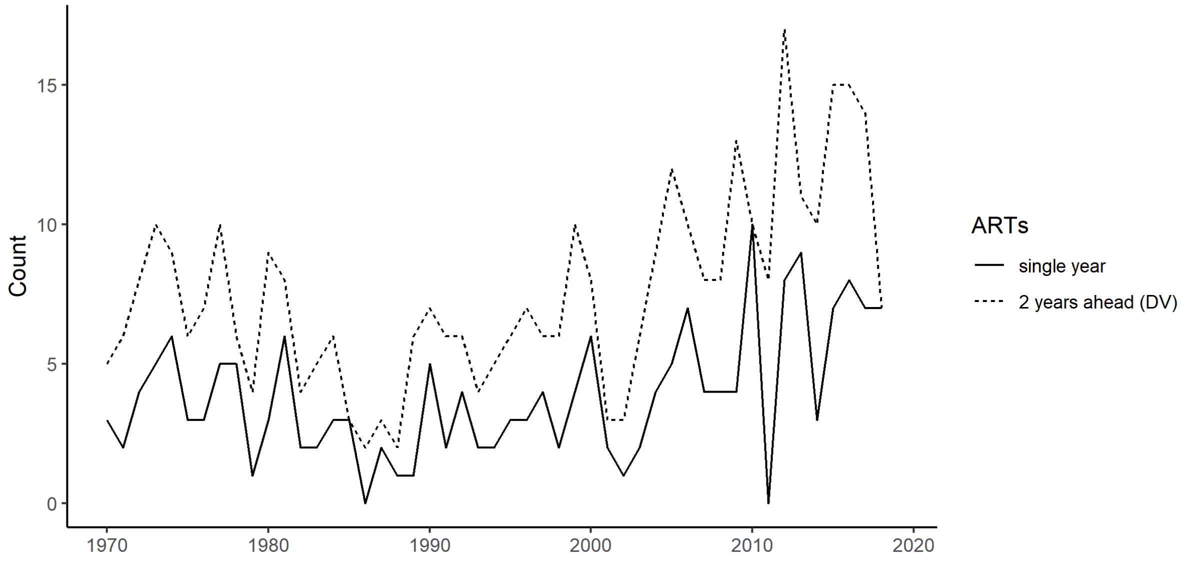
closed autocracy, for example) from one year to the next. We provide a complete list of all ARTs in Appendix D.

One concern with our current operationalization is that we may be identifying small real-world changes when the underlying components of the RoW variable start near the RoW thresholds. In future work, we will assess robustness with alternative operationalizations of ARTs. Nonetheless, we should note that most recorded ARTs represented substantial adverse events or are part of a sustained, but gradual democratic erosion processes. Take, for example, the ART we capture in Togo in 2016 when the regime was downgraded from an electoral democracy to an electoral autocracy. While Togo was a border case on key RoW component variables in 2015, the events that pushed Togo past the threshold were significantly detrimental. For example, in the run-up to the 2015 elections, the government banned all forms of protest and imprisoned political opposition leaders and supporter. Further, Amnesty International reports that a number of those detained were tortured while others were put on trial without access to a lawyer.

Conversely, the ART that we capture in Albania in 2017, for example, seems to be a function of a slight decrease in the Liberal Component index. It was a liberal democracy in 2016 but was downgraded to an electoral democracy in 2017 when its score on the Liberal Component index fell from 0.8 to 0.79, thus below the threshold of 0.8 used by the RoW index to classify a country as a liberal democracy. However, Albania (and similar cases) represent regimes that are influx; they are fragile regimes, unstable in their current class. Our models should be able to identify which of these unstable regimes are in greater need of democracy support.

Table 1 presents the frequency of all movements in the RoW index at yearly level; bold text highlights the frequency of ARTs. There were 189 adverse regime transitions, or roughly 2.4 percent of our country-year observations. There was no change in the RoW index in 93.31 percent of our observations. From 1970 to 2018, the average number of adverse regime transitions per year was 3.86, with a standard deviation of 2.34, and the maximum number of adverse regime transitions in a year was ten in 2010 while two years (1986 and 2011) experienced no adverse regime transitions.

Figure 2: Adverse Regime Transitions by Year



Because we are interested in forecasting the estimated risk that an ART will occur within a *two-year window*, our dependent variable records a country-year observation a one if there is any negative change in the RoW index within the next two years, zero otherwise – e.g., if country X experienced an ART in 2013, this event is reflected in the data (coded as a one) for 2012 and 2013. This increases the number of positive cases to 369 (4.76 percent). This is slightly less than double the number of year-to-year ARTs, as Bolivia, Malaysia, and Somalia experienced an adverse regime transition in the first year of our sample (1970), while other countries experienced adverse regime transitions in two consecutive years (Chile in 1973 and 1974, Argentina in 1976 and 1977, Bangladesh in 2006 and 2007, and Thailand in 2013 and 2014), or had two adverse regime transitions within three years (Cambodia in 1971 and 1973 and Guinea-Bissau in 2008 and 2010).

Figure 2 plots the number of ARTs per year (1970-2018) as well as the number of ARTs that occurred within the two-year window (our dependent variable). The yearly average number of ARTs that occur within a two-year window was 7.53, with a standard deviation of 3.46. The maximum number of adverse regime transitions within a two-year window was 17 in 2012. Two years, 1986 and 1988, experienced the minimum number of ARTs (two) within a two-year window. As the yearly number of countries in our sample ranges from 135 to 169, the effective positive rate of our dependent variable in any given year varies from around 1.4 percent to 10 percent; roughly 75 percent of our yearly positive rates are between 3.8 percent and 5.6 percent.

2.3 Modeling

Looking across the universe of adverse regime transitions, one thing is clear: incumbent political elites can lead their country down a number of different paths that would trigger an adverse regime transition (Coppedge 2017). This equifinality – the different political processes that can lead to an ART – complicates the use of more traditional methodological approaches in prediction problems. Indeed, while traditional methods of description, explanation, and inference can perform well at explaining the onset of important political events *ex-post*, when researchers apply these well-established methods to prediction problems, they tend to perform poorly (Beger et al. 2016, Hill & Jones 2014, Schrodt 2014, Soyer & Hogarth 2012, Ward et al. 2010). To overcome these problems, social scientists have been borrowing forecasting methods from other fields – machine learning methods from computer science, in particular. While machine learning methods are not a panacea (Bowlsby et al. 2019, Cederman & Weidmann 2017), they do offer promise, in the form of improved out-of-sample accuracy, over more traditional methods.

This project uses an unweighted model average ensemble built around three machine learning models: logit with elastic-net regularization, random forest, and gradient boosted forest.[7] Each of these forecasting algorithms takes advantage of the full set of covariates within our expansive dataset, which helps them account for the different, complex, and interrelated political processes that can lead to an adverse regime transition. However, because each of algorithm samples and processes the data differently when calculating the predicted risk of an ART event, some models are better able to predict specific cases. Further, the within method variance of the assigned predicted probabilities tends to be quite different.

For example, relative to our gradient boosted forest model, our random forest model generally produces more gradual, incremental increases in its predicted risk estimates. For our prediction problem, the gradient boosted forest model tends to produce a lot of really low estimates followed by a handful of relatively high estimates. Moreover, the estimates from our logit with elastic-net regularization model tend to be too conservative while our random forest model tends to produce very large estimates. Thus, the unweighted model average ensemble is our preferred approach to estimation, as it not only helps smooth out our predicted risk estimates, but it also improves accuracy, as measured by standard performance statistics.

The machine learning models we use in our ensemble model are becoming increasingly popular among social scientists. For example, the Early Warning Project uses the logit with elastic-net regularization method to forecast mass killings (Woocher et al. 2018), and the random forest model is one of the methods Hegre et al. (2019) at the ViEWS project use to predict various types of political violence. However, while Gohdes (2019) uses the

[7] We provide a brief, nontechnical overview of the mechanics behind each of these machine learning methods in Appendix B.

gradient boosted forest method for a classification problem looking at the relationship between internet accessibility and the type and target of political violence in Syria, to our knowledge, we are the first to adapt this approach to a forecasting problem centered around a distinct political phenomenon.

3 Evaluating model performance

Our framework for evaluating overall model performance is two-fold, and entirely based on out-of-sample performance:

1. Repeated k-fold cross-validation on the training period from 1970 to 2017 in order to obtain out-of-sample predictions for those years.
2. A series of test forecasts for the years from 2011 to 2017 that mimic the process we use to generate our "live" two-year-out forecasts for 2019-2020.

The first approach, the repeated k-fold cross-validation, allows us to assess out-of-sample performance for the complete history the models use for training. The objective of this procedure is to assess a model's ability to predict new, out-of-sample data (i.e., the generalizability of a model, or rather how well a model *should* preform given new data). In short, k-fold k-fold cross-validation (CV) procedures work by randomly dividing all observations into k roughly equally-sized groups. In series, each unique group is set aside to serve as "test data," while the remaining $k - 1$ groups are "training data." These training data are used to fit a model, while the test data is used to evaluate the performance of this fitted model. This process is repeated k times. Thus, each of the k-groups of observations are used once in the testing phase and $k - 1$ times in the training phase. The average performance statistics from these train/test iterations provide researchers with a general sense a model's predictive accuracy.

One concern with the standard k-fold CV approach is that the processes that led to ARTs in the past might be different from the processes that trigger ARTs in more recent years. Thus, it is important to check how well a model performs under conditions that mimic current conditions. The second approach, our set of yearly test forecasts, addresses this concern. These test forecasts give us a better assessment of recent performance and how well the live forecasts are likely to perform. It also gives us some insight in how much performance differs from year to year due to randomness or other fluctuations.

In addition to assessing overall performance, we also examined two additional aspects. The first is calibration, which examines how accurate or valid the probabilities produced by the model are. The second is a simulation experiment where we assess how well the global number of ARTs implied by the model's forecasts matches observed numbers of ARTs.

Table 2: Repeated cross-validation (2×7) performance, 1970-2017

Model	AUC-ROC	AUC-PR	Brier
Logit w/Elastic-net Regularization	0.85	0.30	0.039
Random Forest	0.91	0.40	0.035
Gradient Boosted Forest	0.92	0.42	0.035
Ensemble	**0.93**	**0.46**	**0.034**

We summarize the performance of our models using three standard statistics: Brier, Area Under the Curve-Receiver/Operating Characteristic (AUC-ROC), and Area Under the Curve-Precision/Recall (AUC-PR) scores. The Brier score is the mean squared difference between the predicted probability and the observed outcome. The AUC-ROC metric measures the balance between the *true positive rate* and the *false positive rate* at different acceptance thresholds (the value in which a predicted probability is classified as a one). The AUC-PR metric measures the trade-off between precision, the *positive predictive value*, and recall, the *true positive rate*, across the range of acceptance thresholds.[8]

For the Brier score, the lower the score, the better; however, higher AUC-ROC and AUC-PR scores suggest better model performance. Moreover, an AUC-ROC score above 0.50 signals that the model is preforming better than random chance. And, an AUC-PR score that is higher than the observed frequency of events in the data is a signal that the model is an improvement over random chance. For class-imbalanced prediction problems (when the dependent variable is a rare event) like the one currently under consideration, the AUC-PR measure is better suited since it is more sensitive to how well a model predicts positive cases. Thus, the AUC-PR is our preferred metric of assessment performance.

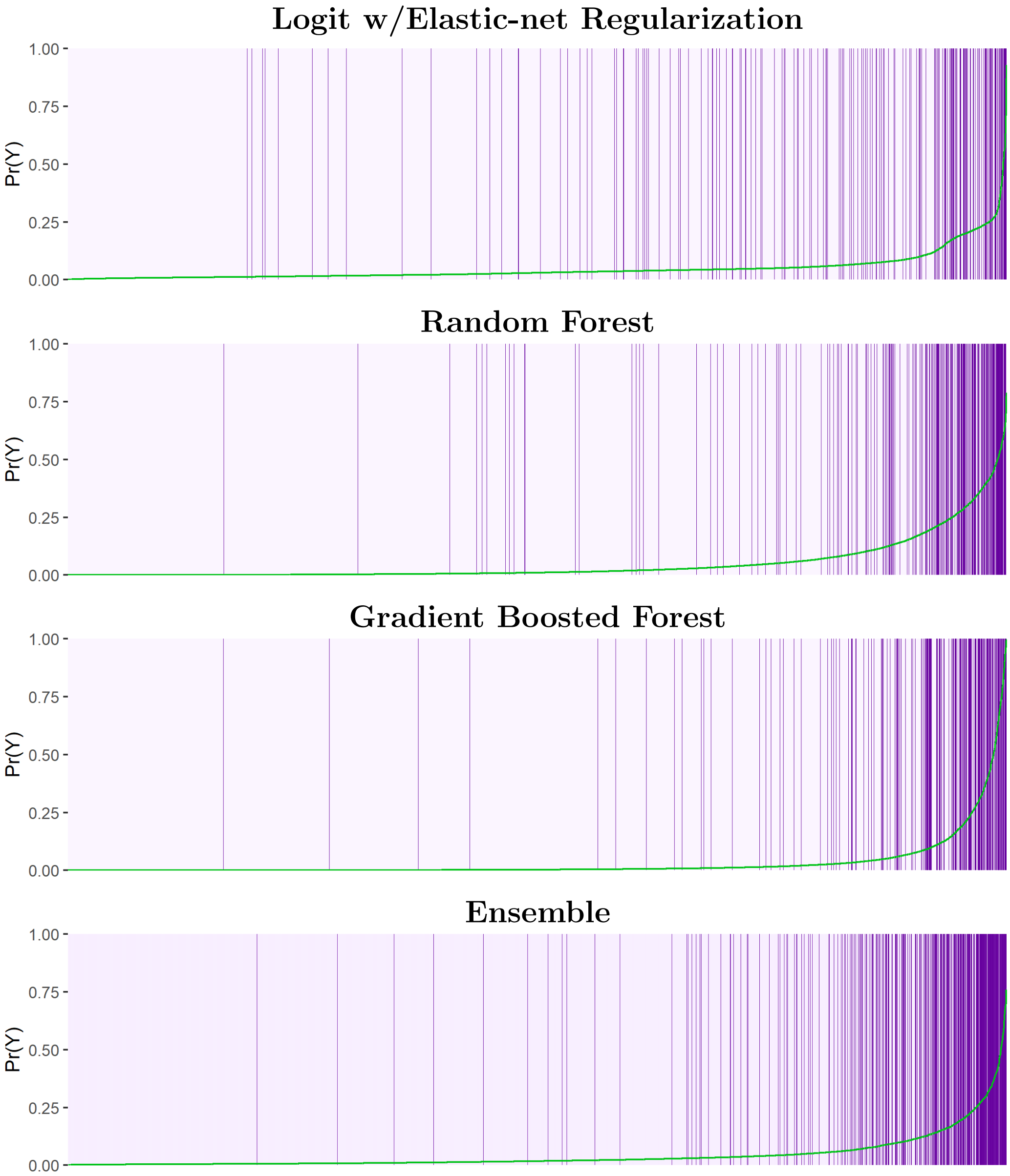
3.1 Cross-validation

Due to the sparsity of our dependent variable, we use a 7-fold CV procedure to help ensure variation in our dependent variable within each test/train split. We repeat this process two times (2×7-fold CV), recording the mean predicted value for each observation and comparing these scores to the observed value. Since this cross-validation assessment covers a larger number of cases, it provides a more robust accuracy estimate.

Table 2 summarizes the results from our 2×7 CV procedure for our ensemble and component models. From Table 2, we see that our ensemble model outperforms our other models across our three performance metrics. Figure 3 presents a series of separation plots based on our 2×7 CV procedure. To produce these plots, we sort the average

[8]True Positive Rate = True Positives / (True Positives + False Negatives); False Positive Rate = False Positives / (False Positives + True Negatives); Positive Predictive Value = True Positives / (True Positives + False Positives).

Figure 3: Separation plots from the 2×7-fold CV procedure, 1970-2017



predicted probability for each observation from our 2×7 CV procedure (the green line running horizontally) and highlight whether an observation had an ART within the two-year prediction window (the dark purple vertical lines). The more concentrated the dark purple lines are to the right of the plot, the better the model performs within our 2×7-fold CV framework.

The green average predicted probability line for each of the three machine learning models highlights the variability of estimates produced by these different methods. The step-like jump at the right side of the elastic-net plot, the relatively higher predictions across the random forest plot, and the sharp, almost exponential increase seen in the gradient boosted forest plot add justification for using an unweighted average model ensemble. As seen in the bottom plot, the ensemble model helps address these differences.

Table 3: Test forecast performance, 2011-2017 (pooled)

Model	AUC-ROC	AUC-PR	Brier
Logit w/Elastic-net Regularization	0.78	0.26	0.064
Random Forest	0.82	0.37	0.059
Gradient Boosted Forest	0.81	0.36	0.061
Ensemble	**0.84**	**0.39**	**0.059**

3.2 Yearly test forecasts

Our series of yearly test forecasts provide a better assessment of the expected accuracy of our "live" forecasts since they more closely replicate those conditions and reflect year-to-year variance in forecast accuracy, which depends heavily on variation in realized adverse regime transitions in a given year. This requires that we conduct a series of train/test experiments. We first train our models using all data from 1970 to 2010. We then use data from 2011 to produce estimated risk forecasts for 2012-13 and evaluate how well our models preformed. We then retrain our models using all data from 1970 to 2011, use data from 2012 to produce estimates for 2013-14, and evaluate model performance. We conduct this iterative model check procedure for all years, 2011 to 2017.

Table 3 summarizes the results for all of the yearly test forecasts. Here too we see that our ensemble model outperforms our other models. Thus, while we present all relevant information for each of the component models in Appendix C, our discussion moving forward will now center on the general performance of our ensemble model.

While pooled summary assessment like those found in Tables 3 and 2 are the standard way to gauge a model's accuracy, the accuracy of the forecasts in any given year is going to deviate to some extent from the overall mean measures. Table 4 show the performance of our ensemble model for each of our yearly test forecasts. For the AUC-ROC column, a score greater than 0.5 is an indicator that the model performed better than the baseline for that year. And, for the AUC-PR a score greater than the observed positive rate in the data, listed under the $\bar{Y}$ column, is an indicator that the model outperformed the baseline for that year.

The AUC-ROC values range from a minimum and maximum of 0.74 to 0.95, with most values in the 0.8 range. The AUC-PR values are more variable, which is expected for rare outcomes like these ARTs. It ranges from 0.15, which is low but still an informative improvement over that year's 0.05 base rate, to 0.51. Overall, there is a lot of variation in performance from year to year; however, the performance metrics for each year suggest that our ensemble model is an improvement over naive models.

Table 4: Performance of the tests forecasts by year, Ensemble model

Year	AUC-ROC	AUC-PR	$\bar{Y}$	ARTs
2011	0.74	0.15	0.05	8
2012	0.88	0.51	0.10	17
2013	0.95	0.41	0.07	11
2014	0.84	0.23	0.06	10
2015	0.81	0.30	0.09	15
2016	0.88	0.49	0.09	15
2017	0.81	0.37	0.08	14

3.3 Model calibration

The model's forecasts are probabilities ranging from zero to one, where values near zero and one indicate near certainty that an event will not or will occur, respectively. In a well-calibrated model, we can take the probabilities at face value, e.g., if we had a set of countries with forecasts of 0.5, we can expect that roughly half of them experience an ART, and we should observe that this was indeed the case in the real world.

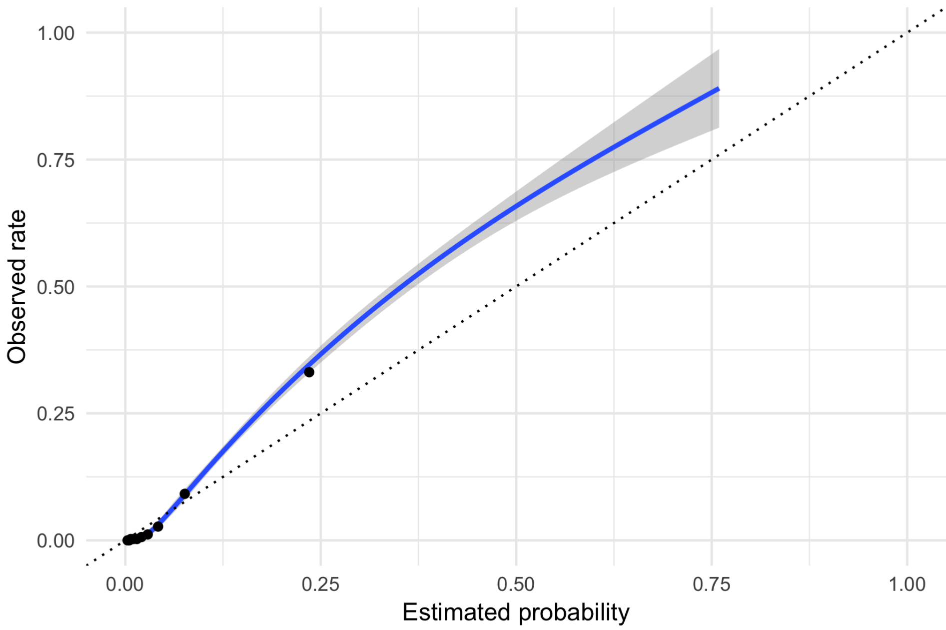
One way to assess this question is to bin the model's probabilities into equally sized groups of similar probabilities, e.g., top ten percent, next ten percent, etc., and compare the number of ARTs implied by the set of probabilities in a bin to the actual number of ARTs that occurred for those countries. Table 5 shows the results if we do this with the full set of cross-validation out-of-sample predictions (i.e., from our 2×7 CV procedure). The first column shows the probability ranges that we use for grouping, or binning, cases. They were chosen to divide the data into 10 equal groups of increasing probabilities. Note that the model's predictions are strongly skewed, with the majority of probabilities falling near zero, and less than a tenth falling above a probability of 0.1. The second and third columns show the average probability for a bin and the implied number of ARTs we would have expected to see. The last two columns show the actual number of ARTs observed and whether, as a result, the model's predictions in that bin were too high ("+") or too low ("−"). The model's forecasted probabilities are slightly too high for lower risk countries, about 80 percent of cases, and too low for the higher risk countries in the remaining 20 percent. Overall, the total number of ARTs predicted out of sample is fairly conservative, with 343 anticipated ARTs against 369 observed ARTs.

Figure 4 is a calibration plot that also shows the model predictions, on the x-axis, against observed outcomes on the y-axis. The black points correspond to the "Mean P" and rate of "Actual ARTs" (i.e., observed ARTs over number of cases in a bin) from Table 5. The blue line is a smoothed estimate of the relationship between the two, which confirms the interpretation above. Below probabilities of 0.05, the model is too aggressive, while above 0.05 it is too conservative and underpredicts.

Table 5: Calibration: expected versus actual number of ARTs by forecast decile, Ensemble model

Bin	Mean P	Expected ARTs	Actual ARTs	Direction
[0.0009, 0.004]	0.003	2.3	0	+
(0.004, 0.006]	0.005	3.9	0	+
(0.006, 0.008]	0.007	5.5	2	+
(0.008, 0.012]	0.010	7.9	2	+
(0.012, 0.017]	0.015	11.5	2	+
(0.017, 0.024]	0.021	16.0	5	+
(0.024, 0.034]	0.029	22.3	9	+
(0.034, 0.053]	0.042	32.5	21	+
(0.053, 0.111]	0.076	59.0	71	−
(0.111, 0.759]	0.235	182.6	257	−

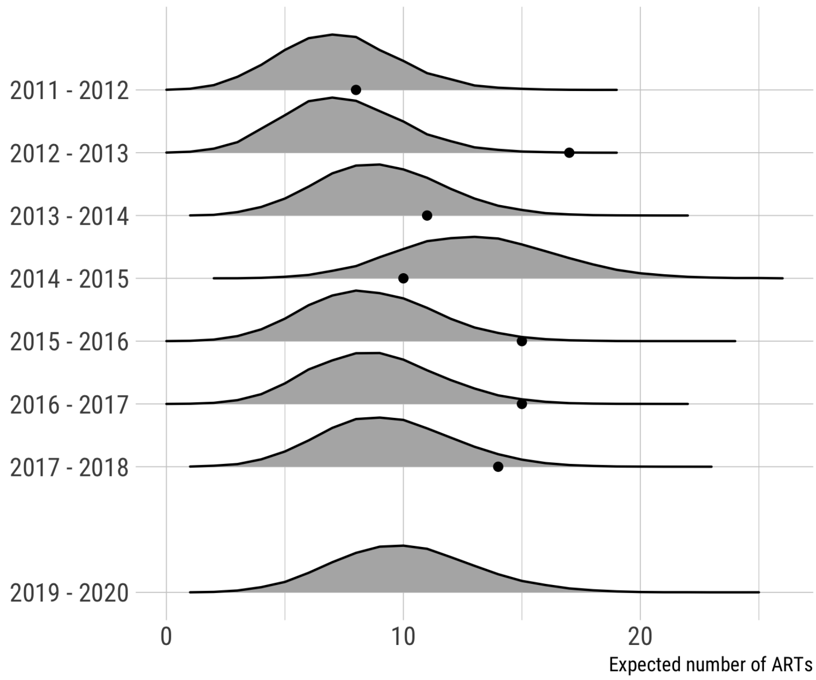
Figure 4: Calibration plot, Ensemble model



3.4 Simulations

Another way to assess the forecasts from our ensemble model is to evaluate how well our yearly test forecast collectively match the real-world outcome we observed in a given year. To do this we can take all of the model's forecasts for a year, which gives us a set of 169 probabilities, and then we "simulate" the world by drawing hypothetical ARTs for each country in accordance with the respective probabilities. It is as if we had a coin to represent each country, flipped them all, and counted the number of heads we get. Except that the coins are all biased in accordance with the model forecasts. With a perfect model, we would expect that the number of hypothetical ARTs we get in this

Figure 5: Expected number of worldwide ARTs given the ensemble model's test forecasts.



simulation to more or less match the number of ARTs we observe in reality. There is a lot of randomness here: one simulation will likely be different from another. So, we do 20,000 of them, and count up how many ARTs each simulated world gave us. We repeat this process for all test year forecasts.

Figure 5 shows the output of our simulation experiment. The gray densities summarize the number of ARTs we got in each of the 20,000 simulated worlds. Each black dot is the actual number of ARTs that happened that within a specific two-year forecast window. So, for example, for the 2011-2012 window, our simulations were pretty close to the observed outcome, whereas for the next 2012-2013 window, the actual number of ARTs was much higher than our simulations would have led us to believe. Overall, the output from these simulations again suggest that our ensemble model is somewhat conservative and under-predicts the global number of ARTs by about two to four in any given forecast window.

4 Current forecasts

Based on the simulations using the forecasts 2019-2020, we expect there to be between five and 15 ARTs overall; 90 percent of our simulations fall into this range. The most likely number of ARTs in our 2019-2020 forecast window is ten; at 13 percent, this is the modal category in our simulation. By doing the simulation exercise with only the top

Table 6: Top 20 Estimated Risks, 2019-2020

	Country Name	Estimated Risk	RoW Classification	Year since last ART
1	Philippines	0.460	Electoral Democracy	14
2	Fiji	0.265	Electoral Democracy	2
3	Mali	0.254	Electoral Democracy	6
4	Hungary	0.250	Electoral Democracy	8
5	Guatemala	0.249	Electoral Democracy	35
6	Kosovo	0.228	Electoral Democracy	6
7	Kyrgyzstan	0.217	Electoral Democracy	2
8	Moldova	0.200	Electoral Democracy	10
9	Bosnia and Herzegovina	0.193	Electoral Democracy	26
10	Tanzania	0.182	Electoral Democracy	2
11	Albania	0.170	Liberal Democracy	1
12	Benin	0.164	Liberal Democracy	3
13	The Gambia	0.143	Electoral Democracy	23
14	Czech Republic	0.129	Liberal Democracy	79
15	Italy	0.126	Liberal Democracy	94
16	Nicaragua	0.122	Electoral Autocracy	11
17	Mauritius	0.120	Liberal Democracy	1
18	Slovenia	0.115	Liberal Democracy	29
19	Turkmenistan	0.106	Electoral Autocracy	6
20	Ghana	0.106	Liberal Democracy	3

20 forecasts, we can also derive some expectations for how many of those countries will experience an ART. This procedure suggests that the most likely number of ARTs in the top 20 is four; 90 percent of our simulations fell between one and seven ARTs.

Table 6 presents these top 20 at-risk countries for 2019-2020. Aside from a country's rank and its estimated risk, Table 6 also includes the country's RoW classification for 2018 and the number of year since it last experienced an ART.[9] For careful observers, a number of these top 20 countries should not come as a surprise. In fact, recent news reports suggest that some of these countries are already displaying signs of adverse regime transition. We discuss a few of these cases below.

The actions of President Duterte of the Philippines over the last few years are reason for concern. Since his election in 2016, President Duterte has demonized journalists, labeling them spies and downplayed the assassination of reporters (Schmidt 2017).[10] Unsurprisingly, Philippines' score on V-Dem's Freedom of expression and alternative sources of information index has declined substantially from 0.90 in 2015 to 0.77 in 2018.

[9] A complete list of all 169 risk forecasts can be found in Table 7 in Appendix A

[10] Schmidt also quotes then-president elect Duterte saying "Just because youre a journalist you are not exempted from assassination if youre a son of a bitch," in 2016 just before his inauguration.

Further, one outcome of President Duterte's so-called war on drugs is a marked increase in extra-judicial killings. This is reflected in a substantial and statistically significant 0.3-point decrease in V-Dem's Physical Violence Index: It went from 0.58 in 2015 to 0.28 in 2018. President Duterte's undermining of democracy include legislative constraints on executive power. President Duterte's governing Coalition for Change enjoys a majority in the Senate (12 out of 24 seats) and a super majority in House of Representatives (258 out of 297 seats) and V-Dem's Legislative constraints on the executive index also decreased substantially and statistically significantly from 0.76 in 2015 to 0.44 in 2018.

The developments in the Philippines in just these last three years are arguably foreboding. Our ensemble model accurately picks up on these as well as a range of other tends in the data, producing a high estimated risk that the Philippines will become an electoral autocracy in the next two years. Indeed, as of this writing, the Philippines have not yet held their mid-term elections, which are scheduled for 13 May 2019. However, some observes fear that these elections will help Duterte further consolidate his power (Kishi & Raleigh 2019)

In Mali, a recent uptick in violence and the government's failure to curb militia violence, led to the resignation of the entire Malian government in April (al Jazeera News Agency 2019, Quashie-Idun & Swails 2019). In Guatemala, President Morales is taking steps to oust supreme court justices who overturned his decision to expel the United Nations-backed International Commission Against Impunity in Guatemala, which has a mandate to root-out corruption high-profile crimes in the country (Amnesty International 2019a, Eulich 2019). Endemic corruption, the erosion of judicial independence, and human rights abuses have pushed people to the streets in protest and the country to a breaking point.

The gradual, yet persistent erosion of democratic norms and institutions in Hungary by Prime Minister Orbán and his far-right Fidesz party has caused alarm bells to ring in Brussels and through a number of European capitals (Foer 2019). Since taking office in 2010, Orbán has worked to curtail judicial independence, erode legislative constrains on the use of executive power, undermine political civil liberties, limit academic freedom, and constrain the ability of civil society organizations to work freely in Hungary. Indeed, the election and immediate actions of Orbán in 2010 pushed Hungary from a promising liberal democracy to an electoral democracy. And, according to V-Dem data, this downward trend has continued unabated for the past ten years.

In Benin, the government recently shut down the internet and cracked down on protest in the run-up to its elections in late April (Amnesty International 2019b, BBC 2019). Further, new election laws in Benin limited the ability of opposition parties to get on the ballet – both of the major parties on the April ballot were loyal the current regime. Moreover, President Talon's government has constrained political civil liberties, such as the right to protest, over the last year. Coupled with the fact that violence erupted after the recent election and that the government is failing to address concerns over the integrity of this election, it seems that Benin is creeping towards an adverse regime transition.

The above descriptions are not to say that any of these cases will, with certainty, experience an adverse regime transition in 2019 or 2020 only that the recent news coming out of these countries is not promising. That said, there are a few countries within the top 20 forecasts seem to be cases in which the current regime is influx; they are bouncing between RoW classes as their political institutions are at the border of different thresholds. As noted in the introduction, Albania is one of these cases; Mauritius and Fiji are others. However, while Albania and Mauritius are examples of a burgeoning liberal democracies that are struggling to solidify important liberalization reforms, Fiji is newly minted electoral democracy with a long history of democratic advancements followed closely by democratic erosion.

Within our sample, Fiji has experienced four ARTs (1987, 2000, 2007, and 2016). V-Dem data suggest that the ART in 2016 was a product of a slight decrease in the Electoral Democracy index. However, the three prior ARTs were all triggered by coups and crackdowns on civil liberties. And, while the military nominally relinquished control in 2015, they handed power over to Prime Minister Bainimarama, the head of the former military regime and leader of the Fiji First party, which is considered to be loyal to the military. In the 2018 elections, the Fiji First party saw its vote share decrease from 59.17 percent to 50.02 percent. Thus, there is concern that if support erodes further the military might again stage a coup.

5 Conclusion

This a paper introduces the V-Forecast project, the Varieties of Democracy Institute's forecasting intuitive. The goal of this project is to develop a suite of forecasting models focused specifically on democratic erosion and other phenomena related to democratic governance. The aim is to make these forecasts easy to understand, transparent in construction, and publicly available. The hope is that policy-makers, aid agencies, and nongovernmental organizations will find these tools useful, and direct resources to at-risk cases. Indeed, we hope that all relevant actors will use these tools to improve the conditions in troubled countries, making our predictions wrong.

Although the current approach and model seem to be working well, two limitations need to be highlighted. First, as the discussion above highlights, our current models perform remarkably well. However, since this is the first iteration of live forecasts where we predict into the actual future (as of the time the forecasts were created), these kinds of accuracy assessments have to be based on retrospective test forecasts. We have to pretend that we do not have information on ARTs and data which in fact we do already have. We did all this with the current, v9 version of V-Dem. The limitation in this, and something that we cannot accurately recreate, is that we do not know whether the underlying data will change as well.

During development of this forecasting system, which used both v8 and later v9 versions of V-Dem, it became apparent that due to an adjustment in some of the data aggregations in V-Dem, and as well the inherent potential for retrospective changes in the estimated V-Dem measurement model, the set of ARTs changed slightly but substantially enough, given the nature of rare events, to negatively impact an attempt to assess v8 model and forecast performance with v9 outcome data. Essentially the target had shifted from under the forecast. This is an unavoidable possibility that we cannot address, but which could impact the future assessment of the current set of forecasts in early 2021.

The other limitation is due to the operationalization of ARTs with the categorical regime of the world indicator. We define an ART as a movement from a higher category to any lower category. Although conceptually ARTs should be significant movements of a regime towards increased authoritarianism, with our current operationalization we cannot distinguish these from instances where a country was already close to the next RoW category experienced a relatively small movement that pushed it across the category boundary. This means that some of the cases of ARTs identified in our data could be relatively small changes in one of the indicators used to construct the RoW categorization.

Moving forward, we plan to expand the scope of this project by developing a series of thematic forecasting models that can estimate the risk that there will be a significant decline in anyone of six dimensions of democracy: (1) elections, (2) civil society, (3) freedom of expression and media, (4) public corruption, (5) democratic inclusion, and (6) legislative and judicial constraints on the use of executive power. Further, we are going to develop a forecasting model that can estimate the risk of state-based electoral violence and intimidation.

By diversifying our forecast targets to include a number of different dimensions of democracy, we can provide policy-makers and aid agencies a suite tools for identifying which specific democratic institutions are at greater risk of erosion, so that they can direct resources accordingly.

References

al Jazeera News Agency. (2019, April 22). *Mali appoints new pm days after government resigns.* https://www.aljazeera.com/news/2019/04/mali-pm-appointed-ogossagou-massacre-190422143413221.html. al Jazeera. (Accessed: April 30, 2019)

Amnesty International. (2019a). *Amnesty international country report: Guatemala 2017/2018.* https://www.amnesty.org/en/countries/americas/guatemala/report-guatemala/. Amnesty International. (Accessed: April 30, 2019)

Amnesty International. (2019b, April 28). *Benin: Internet shutdown on election day is a blunt attack on freedom of expression.* https://www.amnesty.org/en/latest/news/2019/04/benin-internet-shutdown-on-election-day-is-a-blunt-attack/. Amnesty International. (Accessed: April 30, 2019)

BBC. (2019, May 6). *How benin's democratic crown has slipped.* https://www.bbc.com/news/world-africa-48150006. BBC. (Accessed: May 13, 2019)

Beger, A., Dorff, C. L., & Ward, M. D. (2016). Irregular leadership changes in 2014: Forecasts using ensemble, split-population duration models. *International Journal of Forecasting*, *32*(1), 98–111.

Bermeo, N. (2016). On democratic backsliding. *Journal of Democracy*, *27*(1), 5–19.

Bowlsby, D., Chenoweth, E., Hendrix, C., & Moyer, J. D. (2019). The future is a moving target: Predicting political instability. *British Journal of Political Science*, 1–13.

Brandt, P. T., Freeman, J. R., & Schrodt, P. A. (2011). Real time, time series forecasting of inter- and intra-state political conflict. *Conflict Management and Peace Science*, *28*(1), 41–64.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Cederman, L.-E., & Weidmann, N. B. (2017). Predicting armed conflict: Time to adjust our expectations? *Science*, *355*(6324), 474–476.

Coppedge, M. (2017). Eroding regimes: What, where, and when? *V-Dem Working Paper*, *57*.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., ... Ziblatt, D. (2019a). *V-dem codebook v9.*

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., ... Ziblatt, D. (2019b). *V-dem country-year dataset v9.* Varieties of Democracy (V-Dem) Project. doi: https://doi.org/10.23696/vdemcy19

Diamond, L. (2015). Facing up to the democratic recession. *Journal of Democracy*, *26*(1), 141–155.

Eulich, W. (2019, May 10). *Ahead of elections, guatemalas strides and setbacks in justice.* `https://www.csmonitor.com/World/Americas/2019/0510/Ahead-of-elections-Guatemala-s-strides-and-setbacks-in-justice`. The Christian Science Monitor. (Accessed: May 13, 2019)

Foer, F. (2019, June). *Viktor orbn's war on intellect.* `https://www.theatlantic.com/magazine/archive/2019/06/george-soros-viktor-orban-ceu/588070/`. The Atlantic. (Accessed: May 13, 2019)

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, *55*(1), 119–139.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of statistics*, 1189–1232.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research*, *39*(5), 615–637.

Gohdes, A. R. (2019). Repression technology: Internet accessibility and state violence. *American Journal of Political Science*, *Forthcoming*, 1–13.

Goldsmith, B. E., & Butcher, C. (2018). Genocide forecasting: Past accuracy and new forecasts to 2020. *Journal of Genocide Research*, *20*(1), 90–107.

Goldsmith, B. E., Butcher, C. R., Semenovich, D., & Sowmya, A. (2013). Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003. *Journal of Peace Research*, *50*(4), 437–452.

Goldstone, J. A., Bates, R. H., Epstein, D. L., Gurr, T. R., Lustik, M. B., Marshall, M. G., ... Woodward, M. (2010). A global model for forecasting political instability. *American Journal of Political Science*, *54*(1), 190–208.

Hegre, H., Allansson, M., Basedau, M., Colaresi, M., Croicu, M., Fjelde, H., ... Vestby, J. (2019). Views: A political violence early-warning system. *Journal of Peace Research*, *56*(2), 155-174.

Hegre, H., Karlsen, J., Nygård, H. M., Strand, H., & Urdal, H. (2013). Predicting armed conflict, 2010–2050. *International Studies Quarterly*, *57*(2), 250–270.

Hill, D. W., & Jones, Z. M. (2014). An empirical evaluation of explanations for state repression. *American Political Science Review*, *108*(3), 661–687.

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278–282).

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8).

Kishi, R., & Raleigh, C. (2019, May 5). *Midterm elections in the philippines: Power consolidation at its finest.* https://www.acleddata.com/2019/05/10/midterm-elections-in-the-philippines-power-consolidation-at-its-finest/. Armed Conflict Location Event Dataset (ACLED). (Accessed: May 13, 2019)

Kurlantzick, J. (2013). *Democracy in retreat: The revolt of the middle class and the worldwide decline of representative government.* Yale University Press.

Laurent, H., & Rivest, R. L. (1976). Constructing optimal binary decision trees is np-complete. *Information processing letters*, *5*(1), 15–17.

Levitsky, S., & Ziblatt, D. (2018). *How democracies die.* Broadway Books.

Lindberg, S. I., Coppedge, M., Gerring, J., & Teorell, J. (2014). V-dem: A new way to measure democracy. *Journal of Democracy*, *25*(3), 159–169.

Lührmann, A., Dahlum, S., Lindberg, S. I., Maxwell, L., Mechkova, V., Olin, M., ... Stepanova, N. (2018). *V-dem annual democracy report 2018: Democracy for all?* https://www.v-dem.net/media/filer_public/68/51/685150f0-47e1-4d03-97bc-45609c3f158d/v-dem_annual_dem_report_2018.pdf.

Lührmann, A., & Lindberg, S. I. (2019). A third wave of autocratization is here: What is new about it? *Democratization*, 1–19.

Lührmann, A., Mechkova, V., Dahlum, S., Maxwell, L., Olin, M., Petrarca, C. S., ... Lindberg, S. I. (2018). State of the world 2017: Autocratization and exclusion? *Democratization*, *25*(8), 1321–1340.

Lührmann, A., Tannenberg, M., & Lindberg, S. I. (2018). Regimes of the world (row): Opening new avenues for the comparative study of political regimes. *Politics & Governance*, *6*(1).

Mason, L., Baxter, J., Bartlett, P. L., & Frean, M. R. (2000). Boosting algorithms as gradient descent. In *Advances in neural information processing systems* (pp. 512–518).

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2017). How much democratic backsliding? *Journal of Democracy*, *28*(4), 162–169.

Pettersson, T., & Eck, K. (2018). Organized violence, 1989–2017. *Journal of Peace Research*, *55*(4), 535–547.

Powell, J. M., & Thyne, C. L. (2011). Global instances of coups from 1950 to 2010: A new dataset. *Journal of Peace Research*, *48*(2), 249–259.

Quashie-Idun, S., & Swails, B. (2019, April 19). *Mali's prime minister and his entire government resign.* `https://edition.cnn.com/2019/04/19/africa/mali-government-resigns-intl/index.html`. CNN online. (Accessed: April 30, 2019)

Quinlan, J. R. (1979). Discovering rules by induction from large collections of examples. *Expert Systems in the Micro Electronics Age*.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81–106.

Quinlan, J. R. (1987). Simplifying decision trees. *International Journal of Man-Machine Studies*, *27*(3), 221–234.

Rivest, R. L. (1987). Learning decision lists. *Machine Learning*, *2*(3), 229–246.

Rokach, L. (2016). Decision forest: Twenty years of research. *Information Fusion*, *27*, 111–125.

Schmidt, S. (2017, November 14). *Trump chuckled as duterte called journalists 'spies.' that's no joke in the philippines.* `https://www.washingtonpost.com/news/morning-mix/wp/2017/11/14/trump-chuckled-as-duterte-called-journalists-spies-thats-no-laughing-matter-in-the-philippines/`. The Washington Post. (Accessed: April 15, 2019)

Schrodt, P. A. (2014). Seven deadly sins of contemporary quantitative political analysis. *Journal of Peace Research*, *51*(2), 287–300.

Soyer, E., & Hogarth, R. M. (2012). The illusion of predictability: How regression statistics mislead experts. *International Journal of Forecasting*, *28*(3), 695–711.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P., & Girardin, L. (2015). Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family. *Journal of Conflict Resolution*, *59*(7), 1327–1342.

Ward, M. D., & Beger, A. (2017). Lessons from near real-time forecasting of irregular leadership changes. *Journal of Peace Research*, *54*(2), 141–156.

Ward, M. D., Greenhill, B. D., & Bakke, K. M. (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research*, *47*(4), 363–375.

Woocher, L., Zapata, M., Valentino, B., & Hazlett, C. (2018). *Early warning project.* `https://earlywarningproject.ushmm.org/`.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... others (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320.

A Appendix: List of 2019-2020 forecasts for the risk of ART

Table 7: List of complete 2019-2020 ART risk forecasts

	Country Name	Estimated Risk	RoW Classification	Year since last ART
1	Philippines	0.460	Electoral Democracy	14
2	Fiji	0.265	Electoral Democracy	2
3	Mali	0.254	Electoral Democracy	6
4	Hungary	0.250	Electoral Democracy	8
5	Guatemala	0.249	Electoral Democracy	35
6	Kosovo	0.228	Electoral Democracy	6
7	Kyrgyzstan	0.217	Electoral Democracy	2
8	Moldova	0.200	Electoral Democracy	10
9	Bosnia and Herzegovina	0.193	Electoral Democracy	26
10	Tanzania	0.182	Electoral Democracy	2
11	Albania	0.170	Liberal Democracy	1
12	Benin	0.164	Liberal Democracy	3
13	The Gambia	0.143	Electoral Democracy	23
14	Czech Republic	0.129	Liberal Democracy	79
15	Italy	0.126	Liberal Democracy	94
16	Nicaragua	0.122	Electoral Autocracy	11
17	Mauritius	0.120	Liberal Democracy	1
18	Slovenia	0.115	Liberal Democracy	29
19	Turkmenistan	0.106	Electoral Autocracy	6
20	Ghana	0.106	Liberal Democracy	3
21	Comoros	0.104	Electoral Autocracy	3
22	Israel	0.103	Liberal Democracy	8
23	Niger	0.100	Electoral Democracy	8
24	Colombia	0.098	Electoral Democracy	34
25	France	0.096	Liberal Democracy	78
26	Lesotho	0.094	Electoral Democracy	1
27	Haiti	0.092	Electoral Autocracy	13
28	Armenia	0.091	Electoral Autocracy	23
29	Barbados	0.090	Liberal Democracy	118

Table 7: List of complete 2019–2020 ART risk forecasts *(continued)*

	Country Name	Estimated Risk	RoW Classification	Year since last ART
30	Romania	0.086	Electoral Democracy	118
31	Tunisia	0.085	Electoral Democracy	0
32	Somalia	0.085	Closed Autocracy	34
33	Latvia	0.084	Liberal Democracy	29
34	Afghanistan	0.082	Electoral Autocracy	44
35	Chile	0.080	Electoral Democracy	0
36	Burundi	0.077	Electoral Autocracy	22
37	Nepal	0.075	Electoral Democracy	6
38	Pakistan	0.074	Electoral Autocracy	5
39	Guinea-Bissau	0.074	Electoral Autocracy	0
40	Togo	0.074	Electoral Autocracy	1
41	Bhutan	0.073	Liberal Democracy	118
42	Greece	0.071	Electoral Democracy	0
43	Kenya	0.071	Electoral Autocracy	1
44	Lebanon	0.070	Electoral Autocracy	0
45	Macedonia	0.069	Electoral Democracy	6
46	Venezuela	0.069	Electoral Autocracy	12
47	Namibia	0.069	Electoral Democracy	1
48	Central African Republic	0.069	Electoral Autocracy	14
49	Zimbabwe	0.068	Electoral Autocracy	40
50	Costa Rica	0.068	Liberal Democracy	69
51	Timor-Leste	0.067	Electoral Democracy	118
52	Georgia	0.065	Electoral Democracy	28
53	Sudan	0.064	Electoral Autocracy	28
54	Poland	0.063	Electoral Democracy	3
55	Liberia	0.063	Electoral Democracy	14
56	Madagascar	0.063	Electoral Autocracy	8
57	Guinea	0.062	Electoral Autocracy	9
58	Burkina Faso	0.062	Electoral Democracy	3
59	Malaysia	0.062	Electoral Autocracy	48
60	Malawi	0.061	Electoral Democracy	14
61	Uruguay	0.060	Liberal Democracy	45

Table 7: List of complete 2019–2020 ART risk forecasts *(continued)*

	Country Name	Estimated Risk	RoW Classification	Year since last ART
62	Montenegro	0.060	Electoral Autocracy	2
63	South Korea	0.060	Electoral Democracy	0
64	El Salvador	0.056	Electoral Democracy	21
65	India	0.056	Electoral Democracy	43
66	Gabon	0.056	Electoral Autocracy	53
67	Ivory Coast	0.056	Electoral Democracy	4
68	Bolivia	0.055	Electoral Democracy	37
69	Bangladesh	0.055	Electoral Autocracy	6
70	Paraguay	0.054	Electoral Democracy	70
71	Burma/Myanmar	0.054	Electoral Autocracy	56
72	Cape Verde	0.052	Electoral Democracy	0
73	Uganda	0.049	Electoral Autocracy	24
74	Suriname	0.049	Electoral Democracy	38
75	Mozambique	0.047	Electoral Autocracy	44
76	Sri Lanka	0.047	Electoral Democracy	13
77	Nigeria	0.047	Electoral Democracy	15
78	Sierra Leone	0.047	Electoral Democracy	20
79	Turkey	0.047	Electoral Autocracy	5
80	Indonesia	0.046	Electoral Democracy	57
81	Trinidad and Tobago	0.044	Liberal Democracy	118
82	Serbia	0.044	Electoral Autocracy	3
83	Iran	0.043	Electoral Autocracy	118
84	Mauritania	0.042	Electoral Autocracy	10
85	Maldives	0.042	Electoral Autocracy	5
86	Cameroon	0.042	Electoral Autocracy	57
87	Democratic Republic of the Congo	0.041	Electoral Autocracy	53
88	South Africa	0.041	Electoral Democracy	5
89	Panama	0.041	Electoral Democracy	49
90	Botswana	0.041	Electoral Democracy	1
91	Chad	0.040	Electoral Autocracy	49
92	Ecuador	0.040	Electoral Democracy	46
93	Singapore	0.040	Electoral Autocracy	118

Table 7: List of complete 2019–2020 ART risk forecasts *(continued)*

	Country Name	Estimated Risk	RoW Classification	Year since last ART
94	Peru	0.039	Electoral Democracy	26
95	Ukraine	0.039	Electoral Autocracy	6
96	Brazil	0.038	Electoral Democracy	53
97	Azerbaijan	0.038	Electoral Autocracy	28
98	Solomon Islands	0.038	Electoral Democracy	12
99	Guyana	0.038	Electoral Democracy	118
100	Libya	0.038	Closed Autocracy	4
101	Iraq	0.037	Electoral Autocracy	8
102	Croatia	0.037	Electoral Democracy	19
103	Tajikistan	0.037	Electoral Autocracy	28
104	Lithuania	0.037	Electoral Democracy	2
105	Honduras	0.037	Electoral Autocracy	9
106	Ethiopia	0.036	Electoral Autocracy	118
107	Mexico	0.036	Electoral Democracy	118
108	Slovakia	0.036	Electoral Democracy	5
109	Papua New Guinea	0.035	Electoral Autocracy	26
110	Kazakhstan	0.034	Electoral Autocracy	28
111	Dominican Republic	0.034	Electoral Democracy	28
112	Republic of the Congo	0.033	Electoral Autocracy	21
113	Vietnam	0.032	Closed Autocracy	2
114	Zambia	0.032	Electoral Autocracy	3
115	Bulgaria	0.032	Electoral Democracy	69
116	Senegal	0.031	Electoral Democracy	36
117	Cyprus	0.030	Liberal Democracy	49
118	Egypt	0.029	Electoral Autocracy	5
119	Morocco	0.029	Closed Autocracy	118
120	Angola	0.029	Electoral Autocracy	25
121	Finland	0.028	Liberal Democracy	118
122	Mongolia	0.027	Electoral Democracy	107
123	Jamaica	0.026	Electoral Democracy	37
124	Equatorial Guinea	0.025	Electoral Autocracy	39
125	United States of America	0.025	Liberal Democracy	118

Table 7: List of complete 2019–2020 ART risk forecasts *(continued)*

	Country Name	Estimated Risk	RoW Classification	Year since last ART
126	Sweden	0.024	Liberal Democracy	118
127	Austria	0.024	Liberal Democracy	84
128	Taiwan	0.023	Liberal Democracy	8
129	Rwanda	0.023	Electoral Autocracy	44
130	Argentina	0.022	Electoral Democracy	41
131	Cambodia	0.022	Electoral Autocracy	45
132	Belarus	0.022	Electoral Autocracy	21
133	New Zealand	0.021	Liberal Democracy	118
134	South Sudan	0.021	Closed Autocracy	7
135	Canada	0.020	Liberal Democracy	118
136	Djibouti	0.020	Electoral Autocracy	36
137	Portugal	0.020	Liberal Democracy	57
138	Estonia	0.019	Liberal Democracy	26
139	Ireland	0.017	Liberal Democracy	99
140	Russia	0.017	Electoral Autocracy	22
141	United Kingdom	0.016	Liberal Democracy	118
142	Japan	0.015	Liberal Democracy	118
143	Algeria	0.015	Electoral Autocracy	52
144	Thailand	0.013	Closed Autocracy	4
145	Belgium	0.013	Liberal Democracy	78
146	Netherlands	0.012	Liberal Democracy	78
147	Qatar	0.012	Closed Autocracy	118
148	Spain	0.011	Liberal Democracy	79
149	Swaziland	0.011	Closed Autocracy	118
150	Iceland	0.011	Liberal Democracy	88
151	North Korea	0.011	Closed Autocracy	73
152	Germany	0.010	Liberal Democracy	84
153	Jordan	0.010	Closed Autocracy	65
154	Norway	0.010	Liberal Democracy	76
155	Switzerland	0.010	Liberal Democracy	118
156	Denmark	0.010	Liberal Democracy	75
157	Australia	0.009	Liberal Democracy	102

Table 7: List of complete 2019–2020 ART risk forecasts *(continued)*

	Country Name	Estimated Risk	RoW Classification	Year since last ART
158	Uzbekistan	0.009	Closed Autocracy	27
159	Yemen	0.009	Closed Autocracy	2
160	Eritrea	0.007	Closed Autocracy	118
161	Cuba	0.007	Closed Autocracy	58
162	Laos	0.006	Closed Autocracy	27
163	Kuwait	0.006	Closed Autocracy	118
164	Oman	0.006	Closed Autocracy	118
165	Syria	0.006	Closed Autocracy	5
166	Luxembourg	0.006	Liberal Democracy	78
167	China	0.005	Closed Autocracy	97
168	United Arab Emirates	0.004	Closed Autocracy	47
169	Saudi Arabia	0.003	Closed Autocracy	49

B Appendix: Nontechnical overview of machine learning methods

For our unweighted model average ensemble, we use the estimates from three machine learning models: logit with elastic-net regularization, random forest, and gradient boosted forest. As their names imply, logit with elastic-net regularization models are built around maximum-likelihood principles, while random forest and gradient boosted forest start with decision trees. Below, we provide an overview of these machines learning methods.

Logit w/elastic-net regulation: Like standard logistic regression, the elastic-net regularization (ENR) version estimates coefficients for a linear equation relating input covariates to the binary outcome variable through a logistic function. In addition to the regular logistic likelihood, the cost function for elastic net regression includes a penalty term for non-zero coefficient values. This penalty term is governed by two hyperparameters and has the practical effect of pushing some coefficient estimates completely to zero, and shrinking the remaining coefficient estimates towards zero.

The ENR model works well with variables that are highly correlated (provides different weights to highly correlated variables according to how much variation the variable explains, allowing uninformative variables to go to zero), helps reduce model-fit (introduces bias to the training model in order to produce more accurate out-of-sample predictions), and provides a heuristic for variable selection (orders variables according to post-weighted parameter estimates and identifies which parameter went to zero). Introduced by Zou & Hastie (2005), this algorithm has become a mainstay in the machine learning discipline. Within the social sciences, the Early Warning Project uses this method to forecast mass killings (Woocher et al. 2018).

Random Forest & Gradient Boosted Forest: Decision trees are at the root of both the random forest and gradient boosted forest algorithms. Decision tree algorithms start by measuring how well each variable, at different cut-points, classifies observations according to the outcome variable, selecting the cut-point that performs best. The algorithm does this for all variables and determines which variable, and at what cut-point, best explains the outcome variable. This variable becomes a *root node*. The algorithm splits all of the observations according to the root node cut-point, branching out to create two (or more) *sub-nodes*. At each sub-node, the algorithm again assesses which of the remaining variables, and at what value, best classifies the data in each sub-node. It repeats this process to the point in which new nodes (splits) no longer improve classification.

With their introduction and formalization by Quinlan (1979), decision tree-based approaches serve as the foundation of many modern machine learning algorithms (Laurent & Rivest 1976, Quinlan 1986, 1987, Rivest 1987, Rokach 2016). However, while decision trees do well at describing the data at hand, overfitting reduces their out-of-sample accuracy.

First introduced by Ho (1995, 1998) and extended by Breiman (2001), the random forest (RF) algorithm helps address this overfitting issue. To do so, the RF algorithm introduces two randomization techniques. First, it draws a bootstrapped (with replacement) dataset and grows a decision tree. However, rather than assessing how well *all* variables perform at *each* node, the RF algorithm randomly selects m number of variables for consideration. It repeats this process hundreds of times, growing a forest of diverse decision trees, each with a unique set of variables at each node. The observations excluded from the bootstrapped sample are fed into each randomized tree for classification. The algorithm calculates the average classification from these trees and records it as the probability the observation belongs in a specific class, whether there was an adverse regime transition, for example. These predictions are then compared to the observed value of the outcome variable.

Introduced by Breiman (1996), this bootstrap aggregation process (also known as "bagging") reduces issues related to model fit as well as a model's variance – how far the predicted value deviates from the observed value. This out-of-sample error (the mean prediction error) is an important metric for model tuning – determining the optimum number of variables, m, to use as well as the number of trees to grow. Since its introduction, the RF method has quickly become one of the most widely used machine learning algorithms (Rokach 2016). This is one of the primary methods Hegre et al. (2019) at the

ViEWS project use to predict various types of political violence.

Like the RF algorithm, the gradient boosted forest (GBF) algorithm first draws a bootstrapped sample (with replacement) and grows a decision tree. Unlike the RF method that builds and combines a forest of randomly different trees in parallel, the GBF algorithm builds a *series* of trees, where each successive tree is trained so that it attempts to reduce the predictive error of the previous trees. Further, while the RF algorithm grows trees using m number of variables at each node, the GBF method uses the entire variable set at each node split. However, it restricts the number of nodes (splits) within each decision tree, usually between one and ten nodes, growing what researchers refer to as "weak learners" or "shallow trees." The algorithm calculates the residuals using the bootstrapped data. It then fits another shallow tree *to these residuals*. The algorithm combines the two decision trees, runs the observed data through this ensemble of trees, calculates the residuals, and fits a new tree to these residuals. It then combines this tree with the others, repeating this recursive learning process hundreds of times, growing a forest of *dependent* decision trees. Thus, the GBF method has four parameters that affect the model's performance: the number of trees to grow, the number of nodes to allow, and two parameters that regulate how much each successive tree should "learn" (Rokach 2016).

The gradient boosted method was first introduced in the late 1990's and early 2000's as a way to improve the predictive (and classification) power of other decision tree models (Freund & Schapire 1997, Friedman 2001, 2002, Mason et al. 2000, Rokach 2016). This method has grown in use throughout other disciplines; in fact, Wu et al. (2008) regard model boosting algorithms to be one of the top-ten tools for practical machine learning problems. Nevertheless, while Gohdes (2019) uses the GBF method for a classification problem looking at the relationship between internet accessibility and the type and target of political violence in Syria, to our knowledge, we are the first to adapt this approach to a forecasting problem centered around distinct political phenomenon.

Tuning of Model Hyperparameters: We optimized the hyperparameters for all of our machine learning models to maximize out-of-sample fit, with estimates obtained through a model tuning cross-validation procedure. Note that this cross-validation is nested within the top-level 2×7-fold cross-validation scheme we use to assess out-of-sample fit and model evaluation. In short, for each iteration of the top-level cross-validation, we take the current training data set and perform an additional round of model-level cross-validation in order to optimize the hyperparameters.

C Appendix: Performance for ensemble sub-models

C.1 Yearly test forecasts

The plots below help provide a sense of how well each of our models are preforming across a series of yearly test forecasts (2011-17). Throughout these plots, light purple denotes country-year observations without an ART, blue denotes an observation that experienced an ART in the first year of the two-year window, while dark purple denotes observations with an ART in the last year of the two-year window. For example, the figure below shows the risk estimates our ensemble model generated for 2011-12. After training the model with data from 1970 to 2009, we use data from 2010 to calculate the country-level risk of ARTs for 2011-12.

Figure 6: Ensemble: 2011-2012 Test Forecast

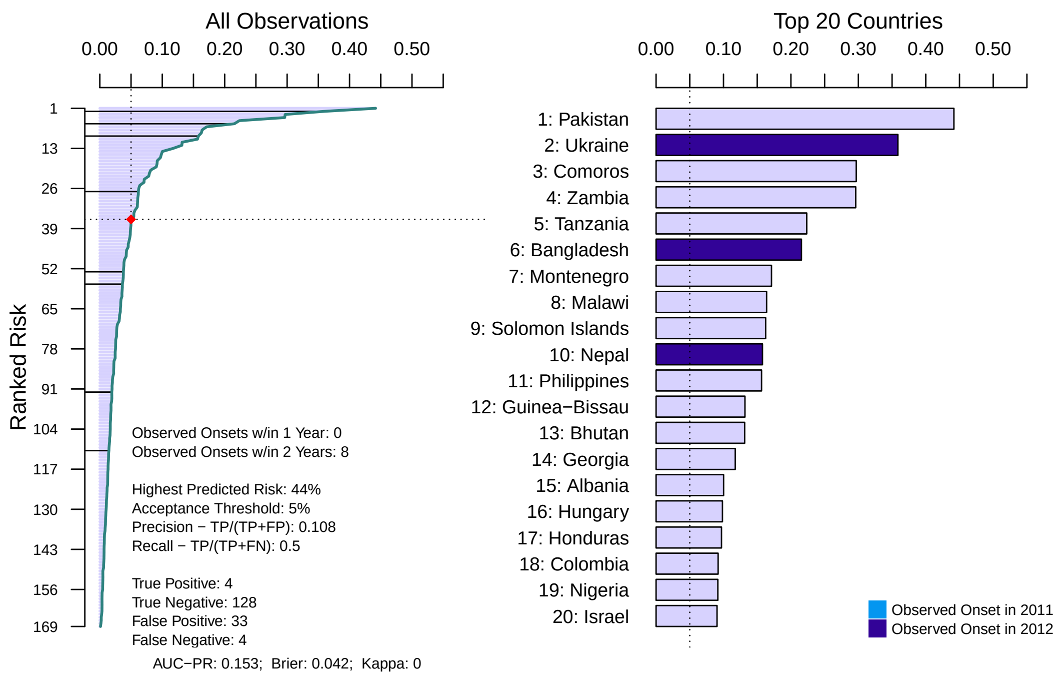


Figure 7: Ensemble: 2012-2013 Test Forecast

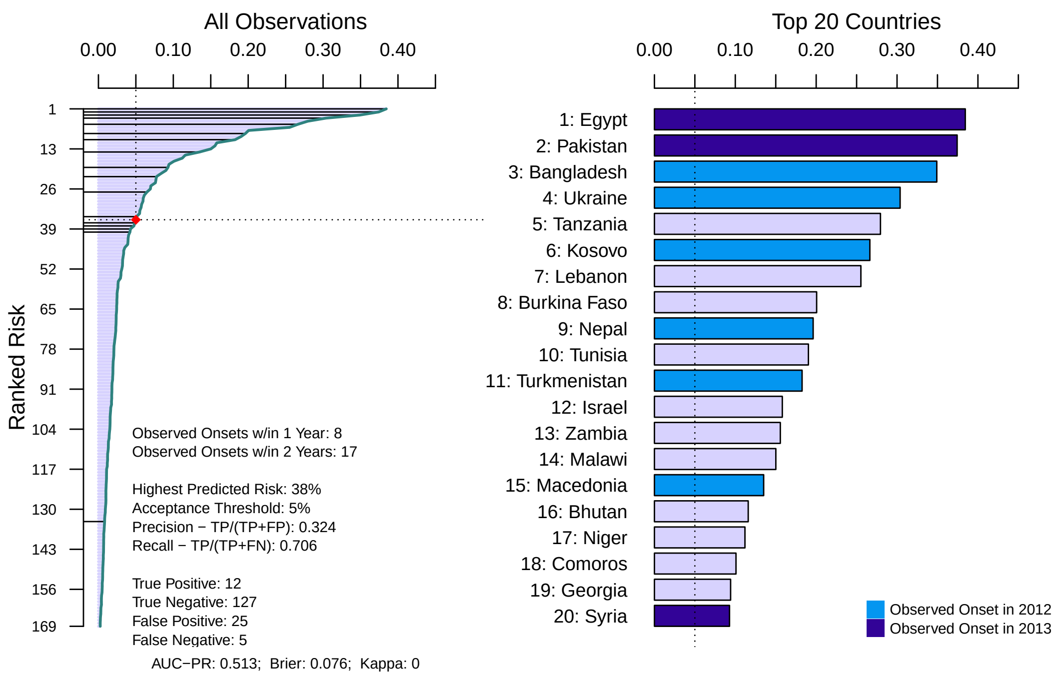


Figure 8: Ensemble: 2013-2014 Test Forecast

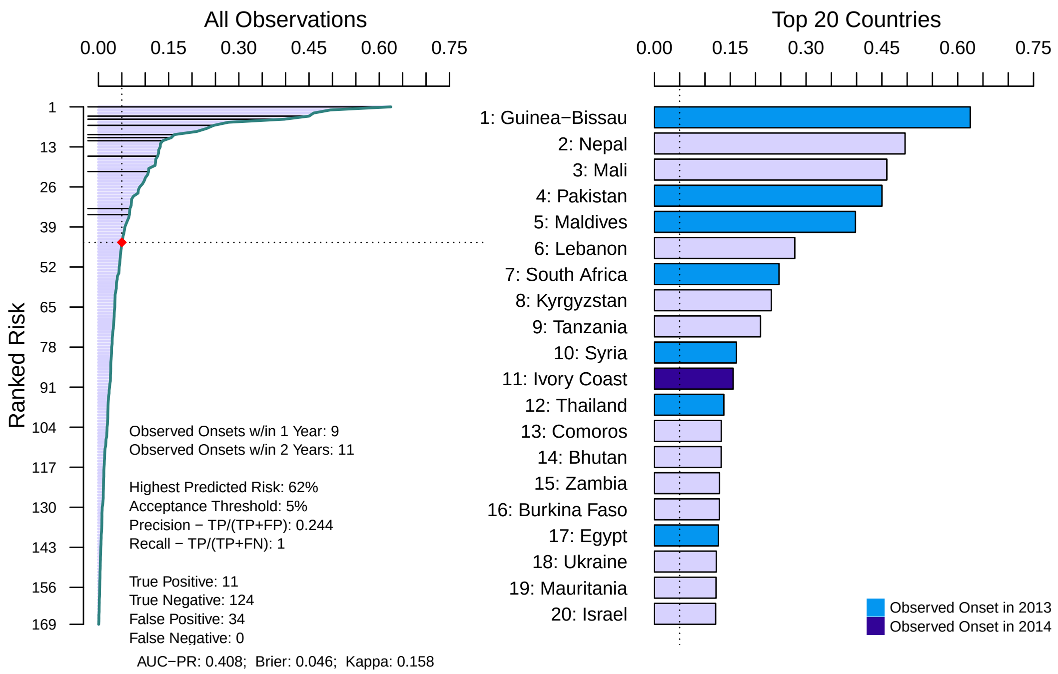


Figure 9: Ensemble: 2014-2015 Test Forecast

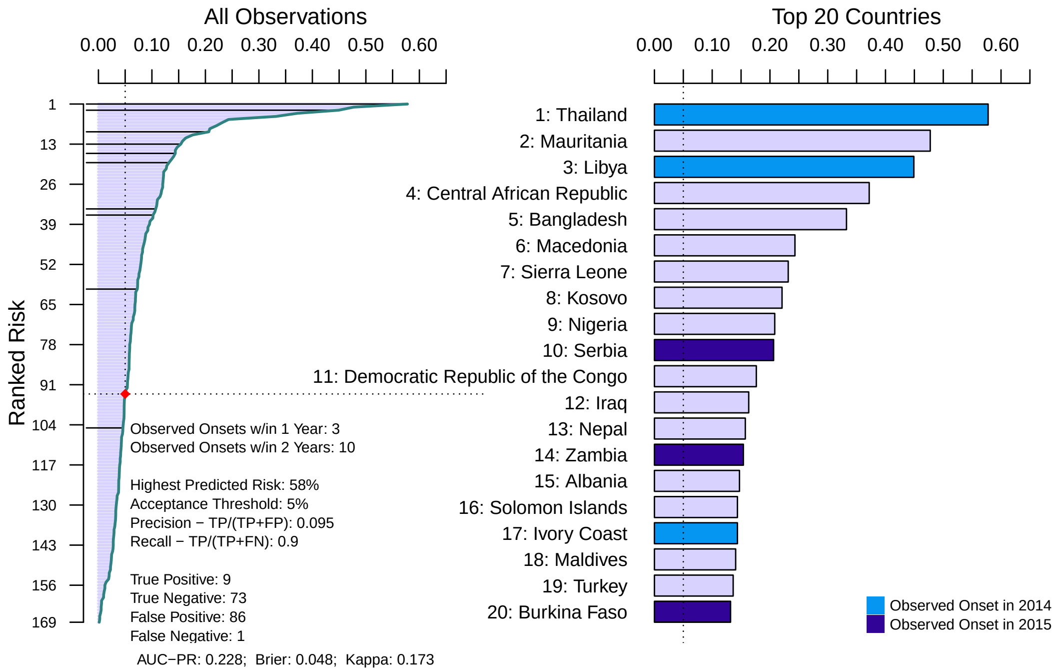


Figure 10: Ensemble: 2015-2016 Test Forecast

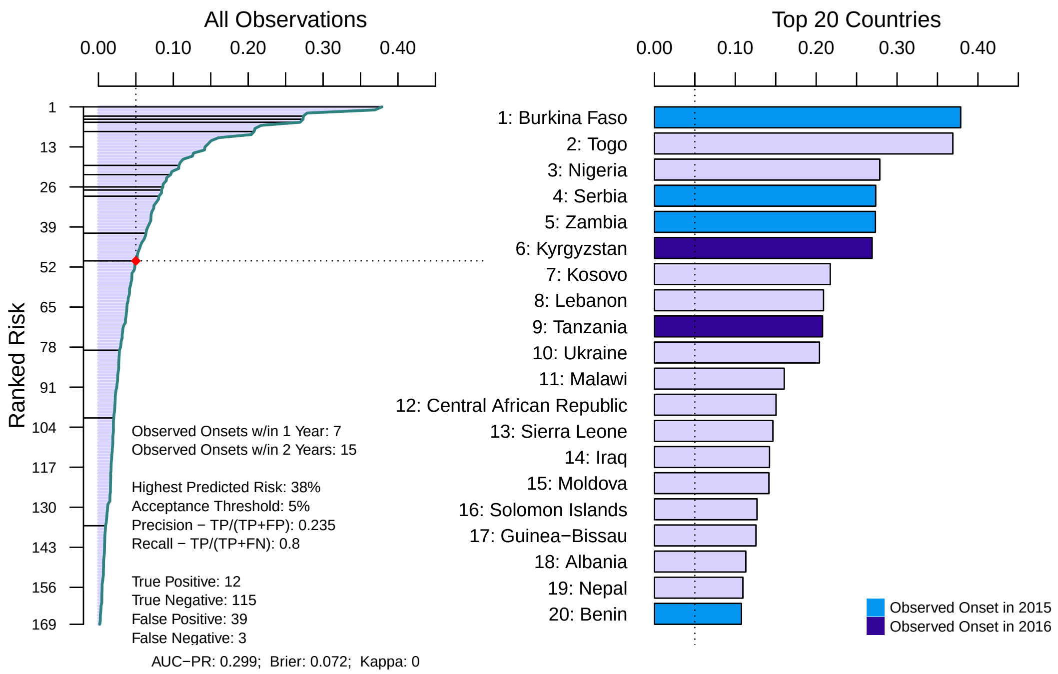


Figure 11: Ensemble: 2016-2017 Test Forecast

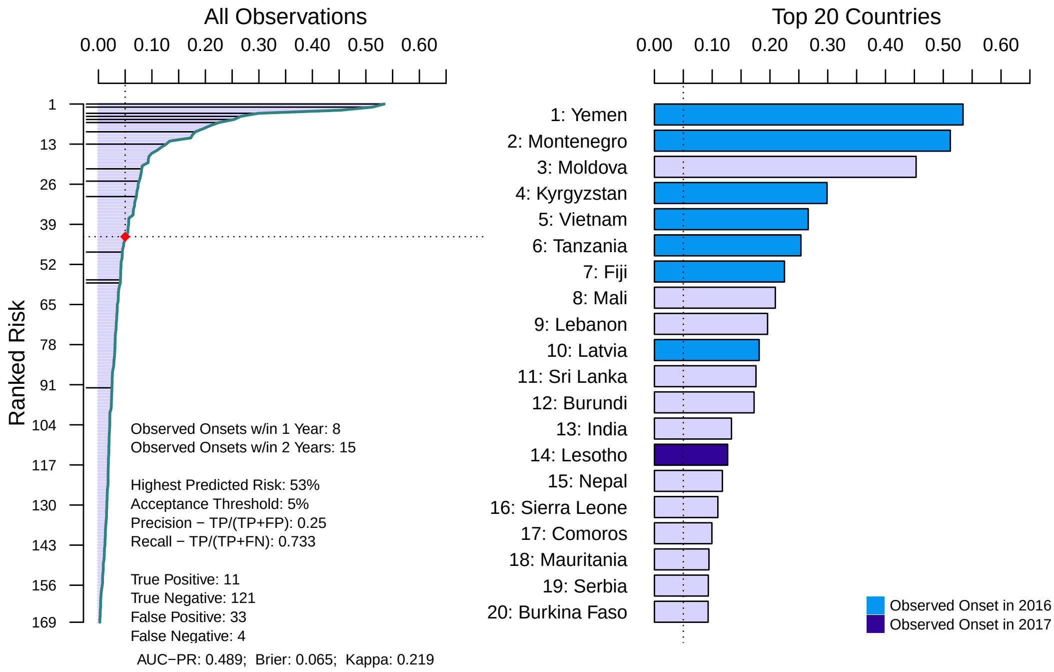


Figure 12: Ensemble: 2017-2018 Test Forecast

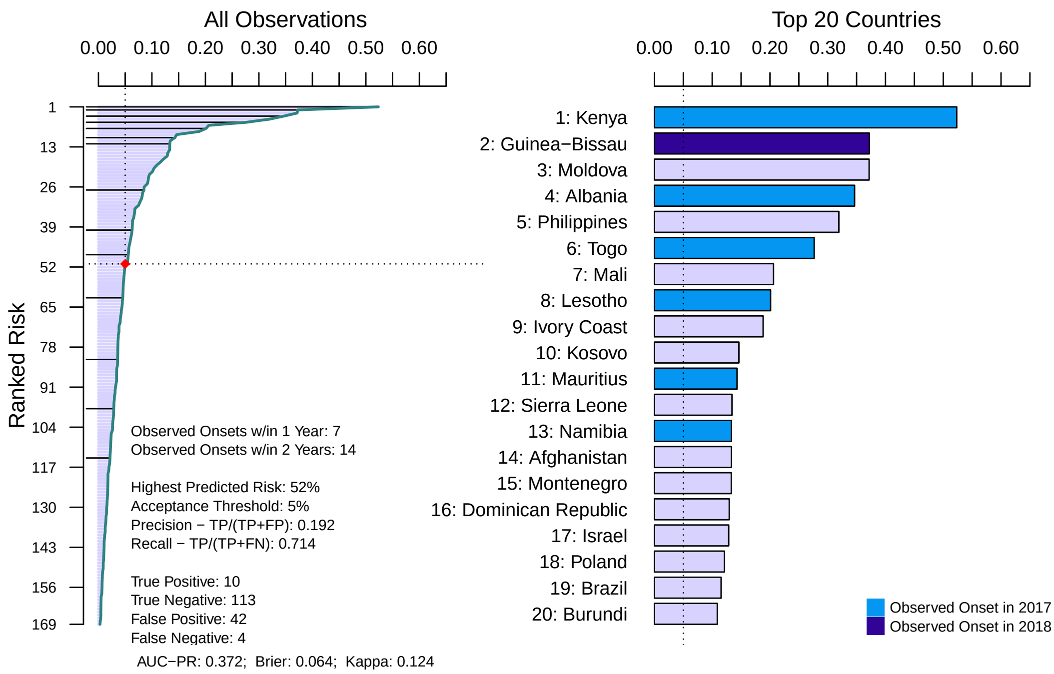


Figure 13: Ensemble: 2019-2020 Forecast

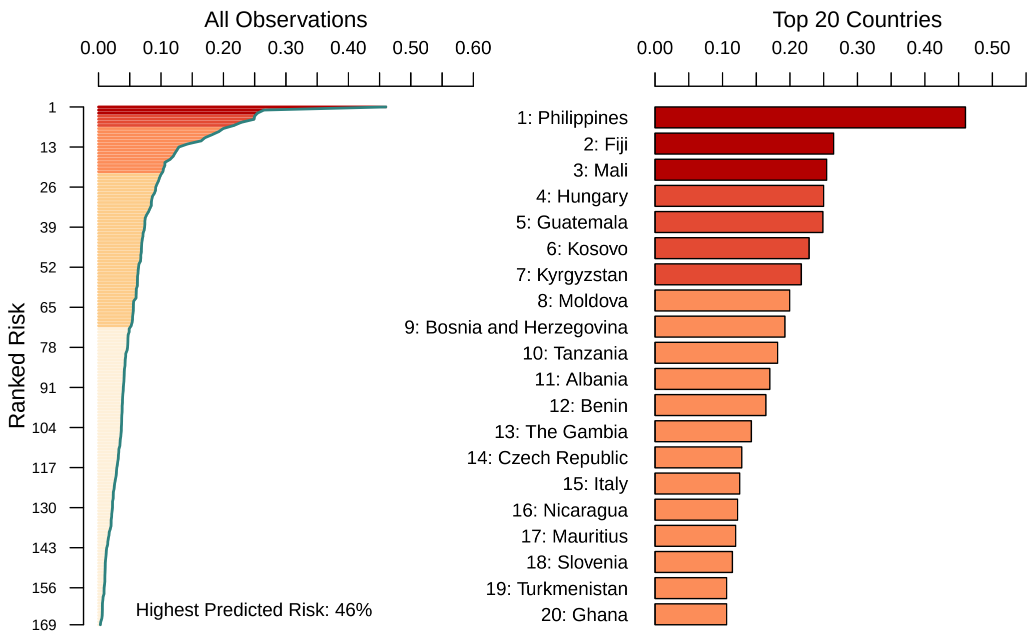


Figure 14: Logit w/elastic-net regularization: 2011-2012 Test Forecast

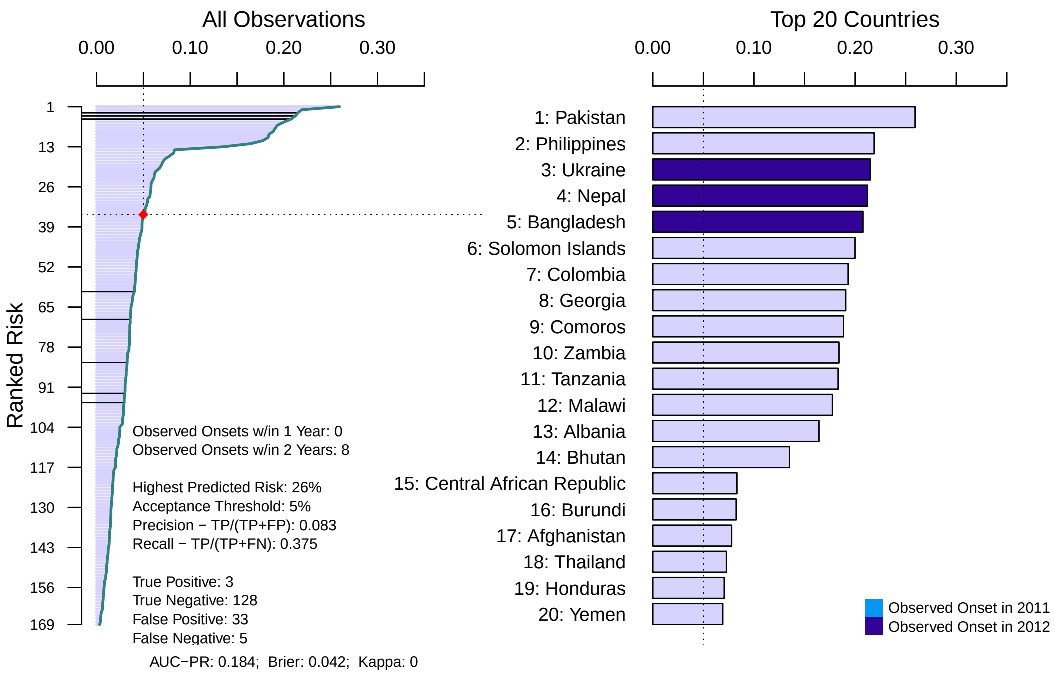


Figure 15: Logit w/elastic-net regularization: 2012-2013 Test Forecast

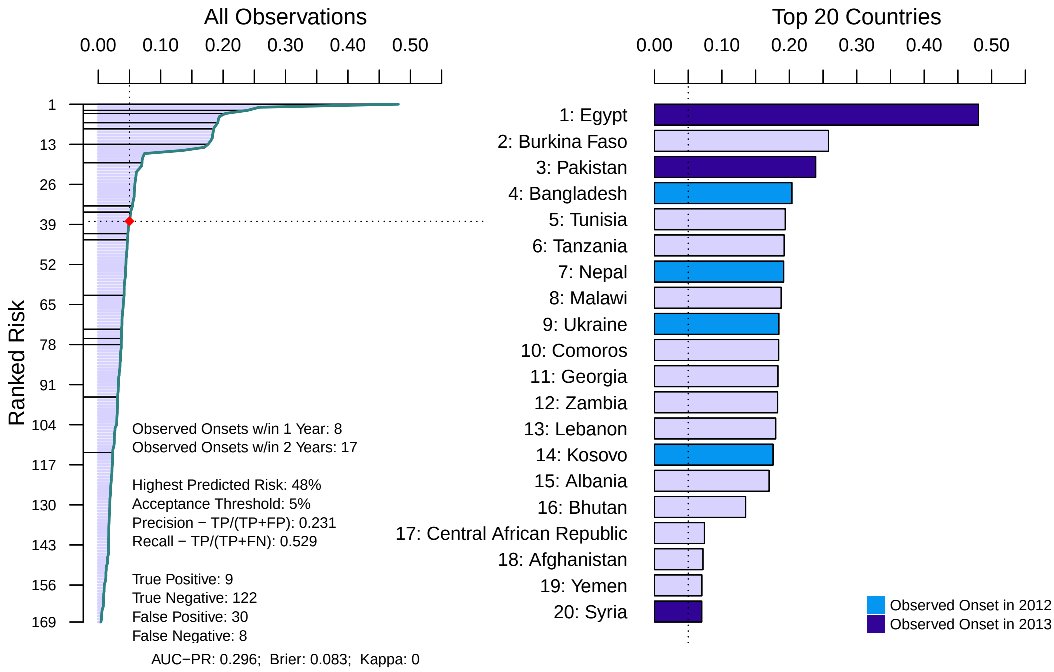


Figure 16: Logit w/elastic-net regularization: 2013-2014 Test Forecast

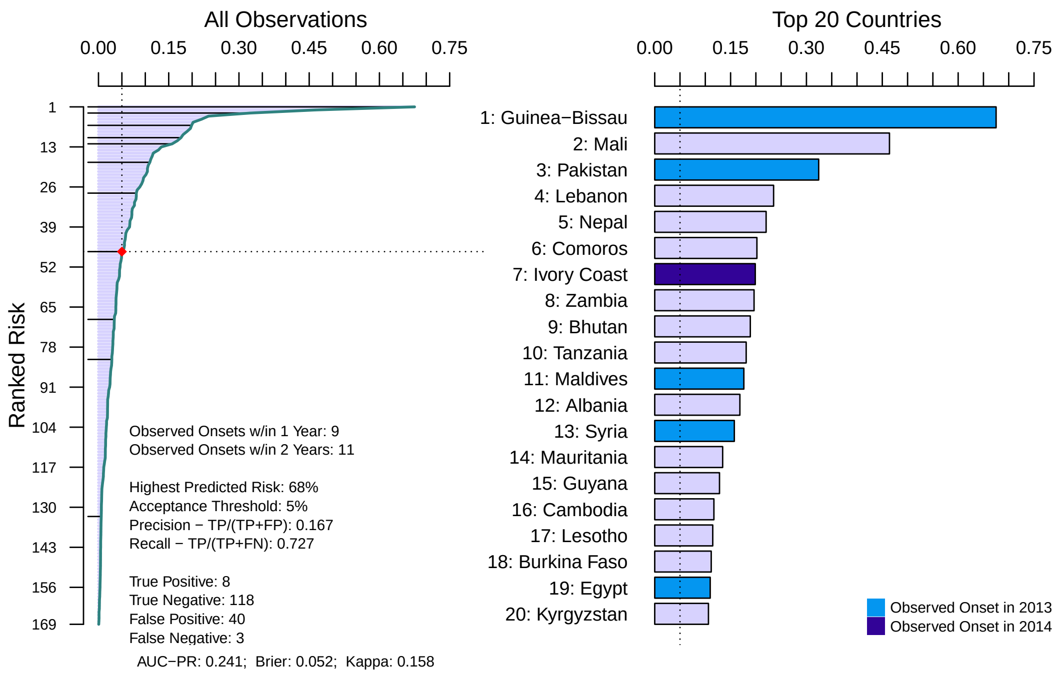


Figure 17: Logit w/elastic-net regularization: 2014-2015 Test Forecast

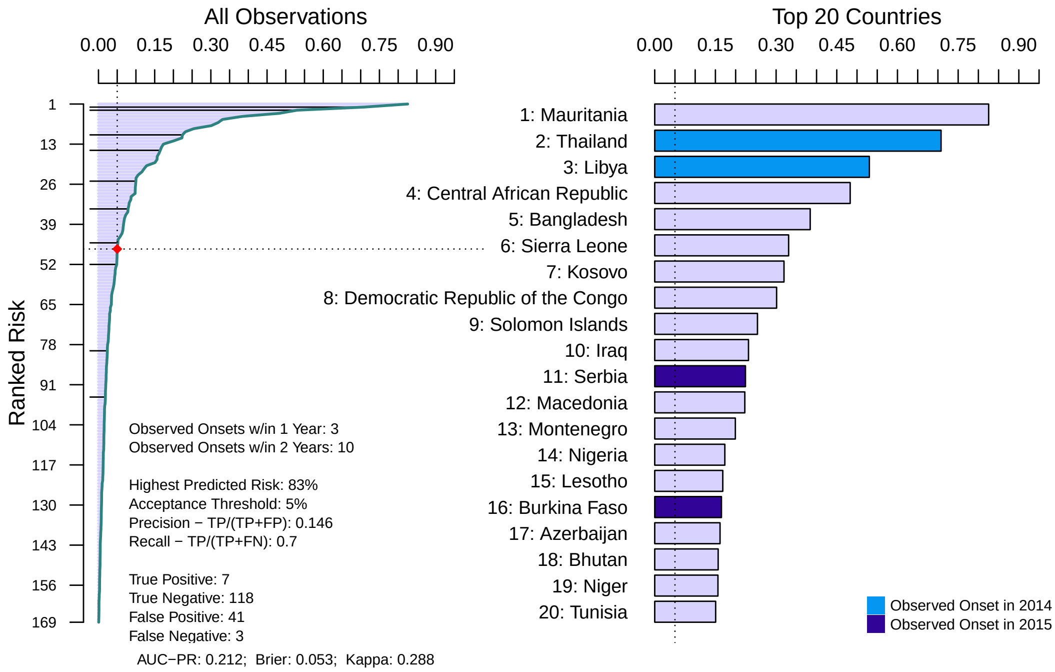


Figure 18: Logit w/elastic-net regularization: 2015-2016 Test Forecast

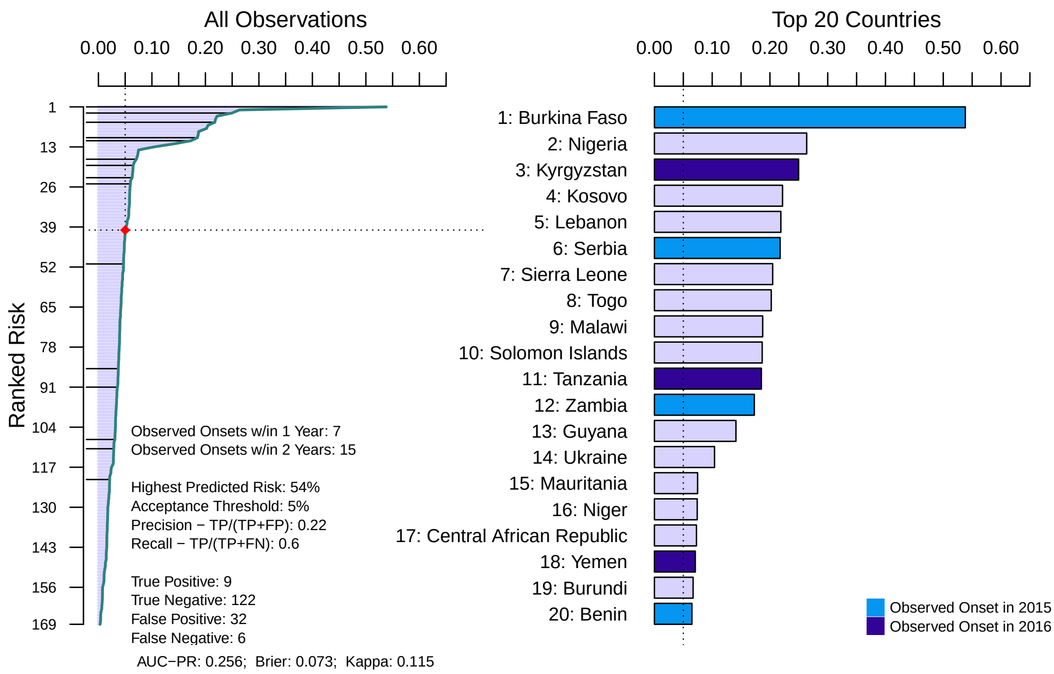


Figure 19: Logit w/elastic-net regularization: 2016-2017 Test Forecast

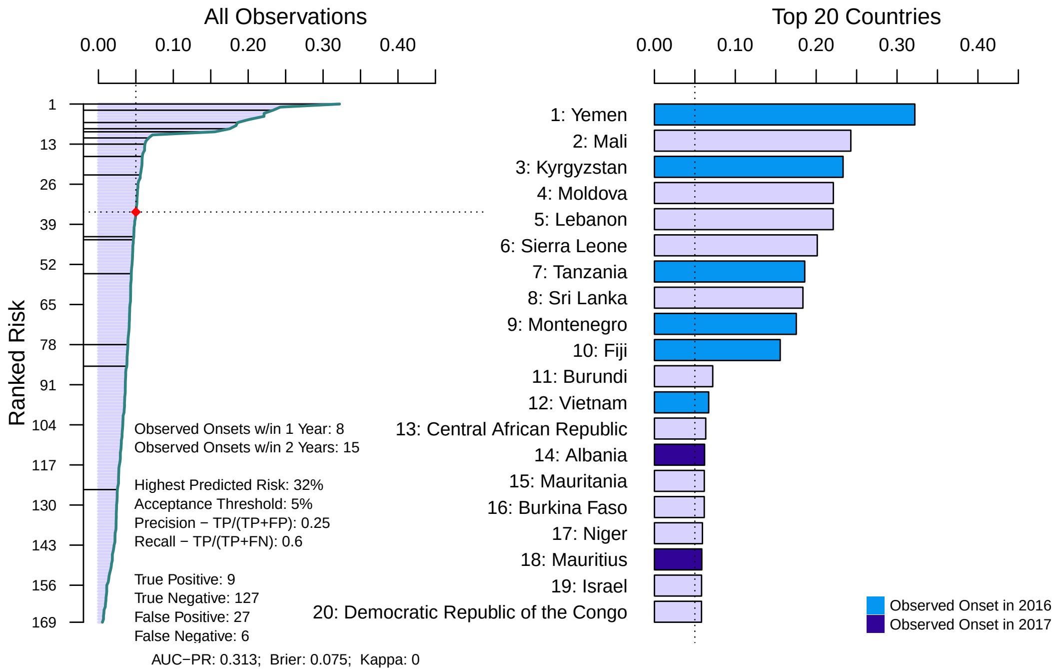


Figure 20: Logit w/elastic-net regularization: 2017-2018 Test Forecast

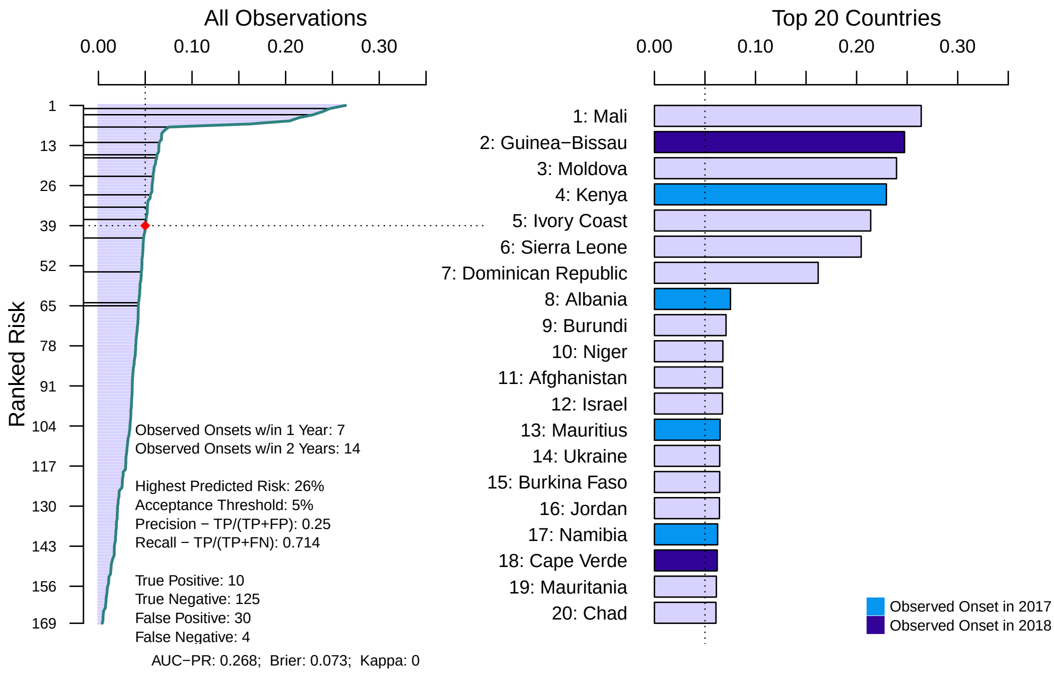


Figure 21: Logit w/elastic-net regularization: 2019-2020 Forecast

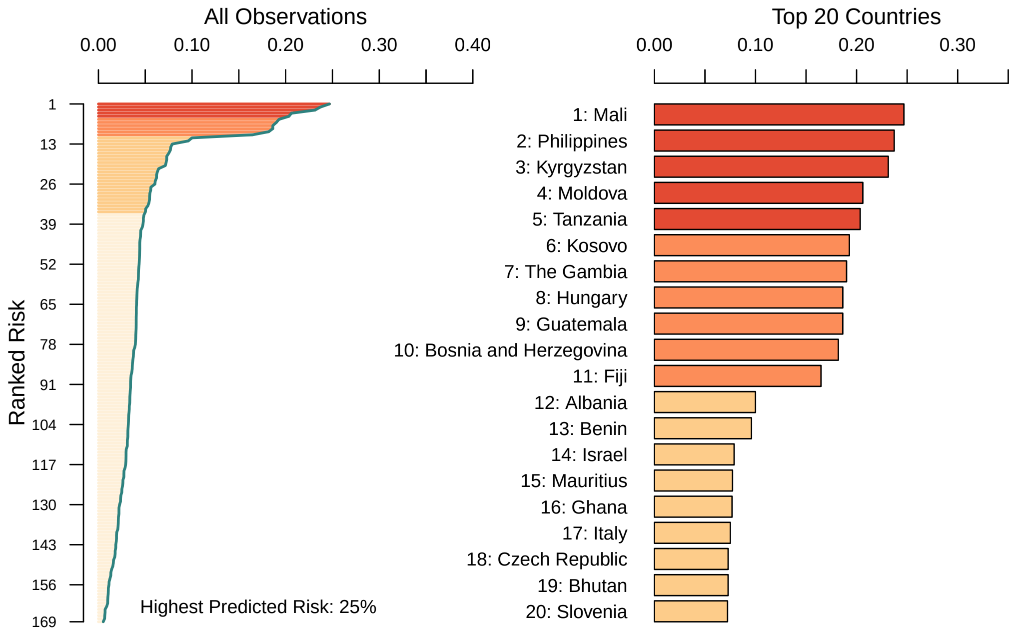


Figure 22: Random forest: 2011-2012 Test Forecast

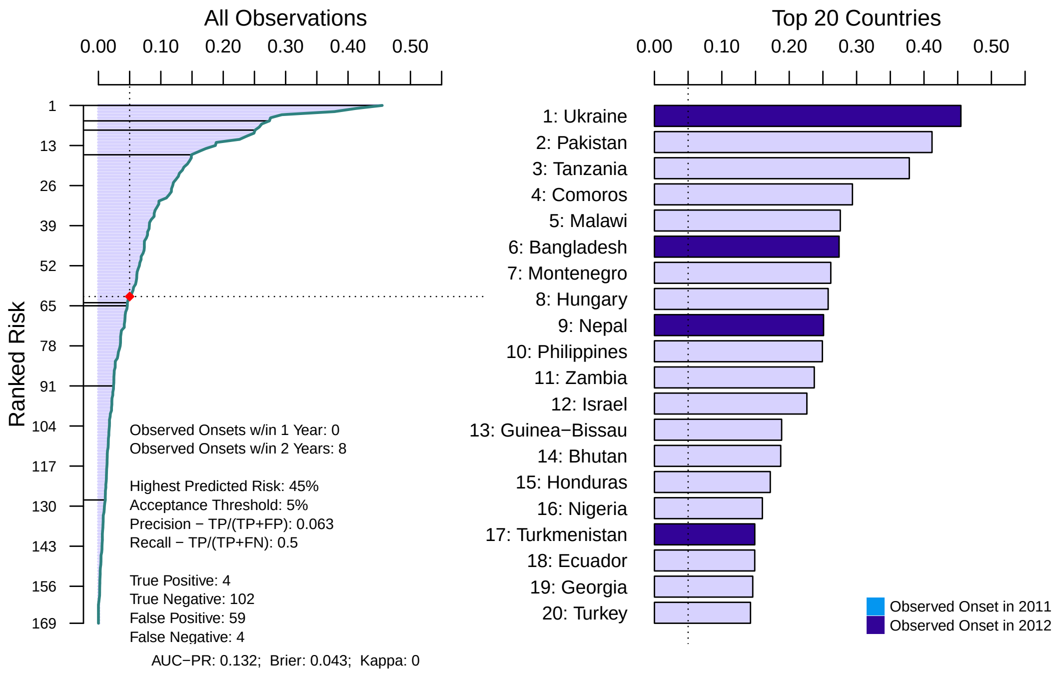


Figure 23: Random forest: 2012-2013 Test Forecast

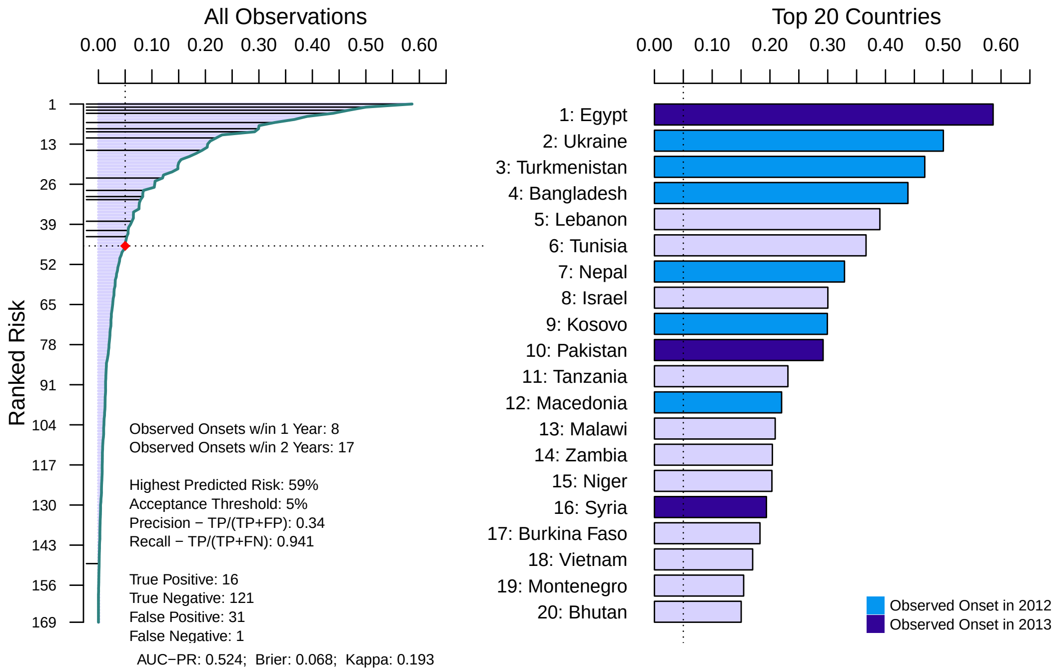


Figure 24: Random forest: 2013-2014 Test Forecast

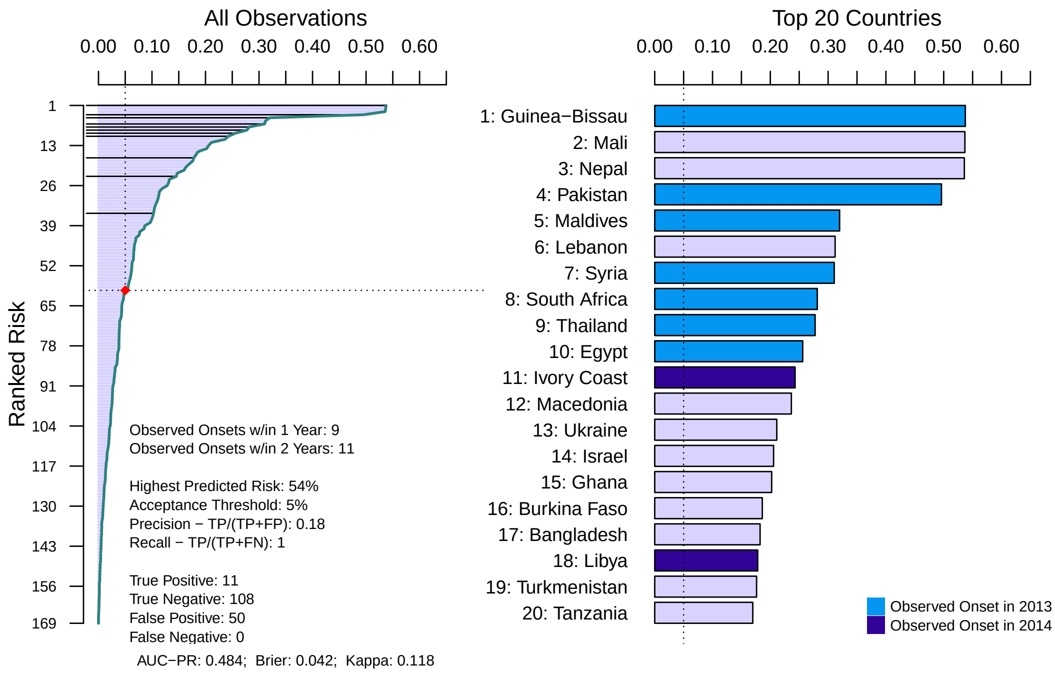


Figure 25: Random forest: 2014-2015 Test Forecast

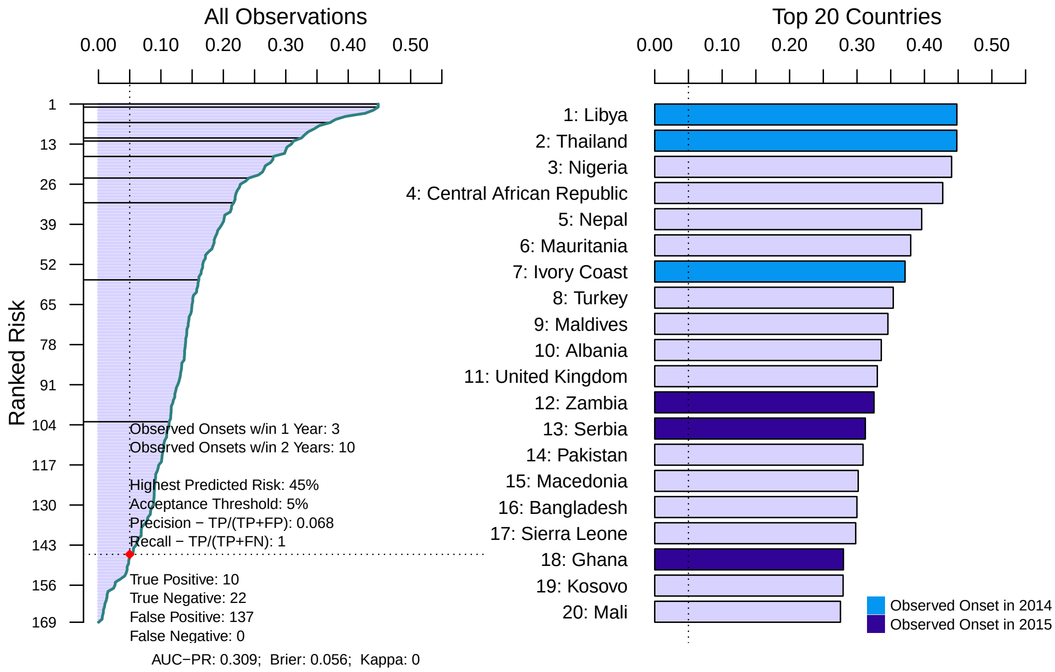


Figure 26: Random forest: 2015-2016 Test Forecast

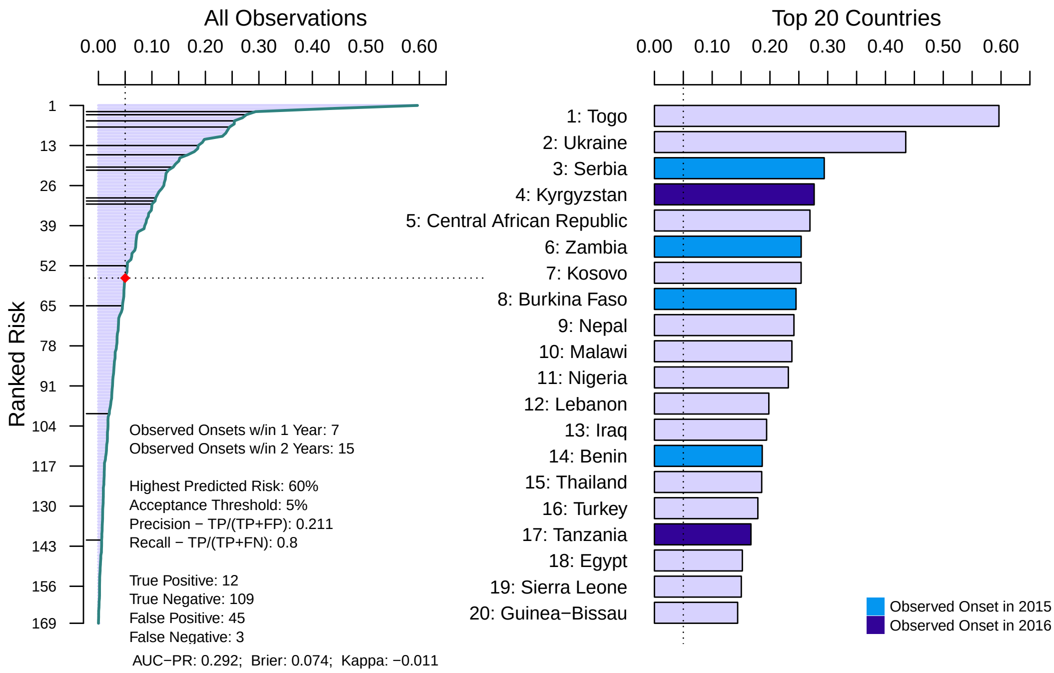


Figure 27: Random forest: 2016-2017 Test Forecast

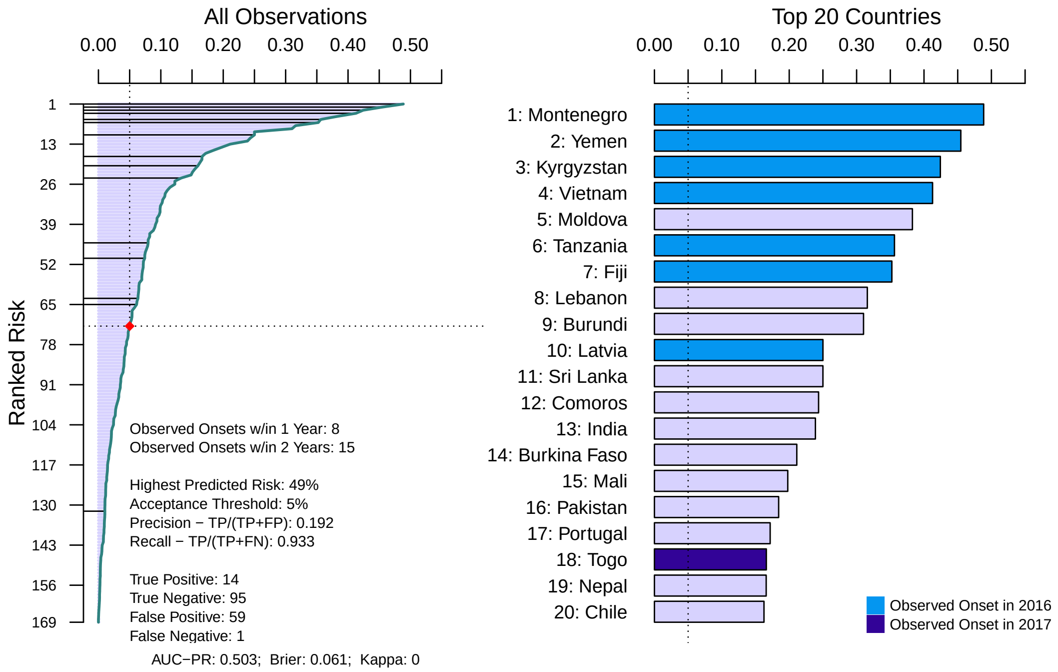


Figure 28: Random forest: 2017-2018 Test Forecast

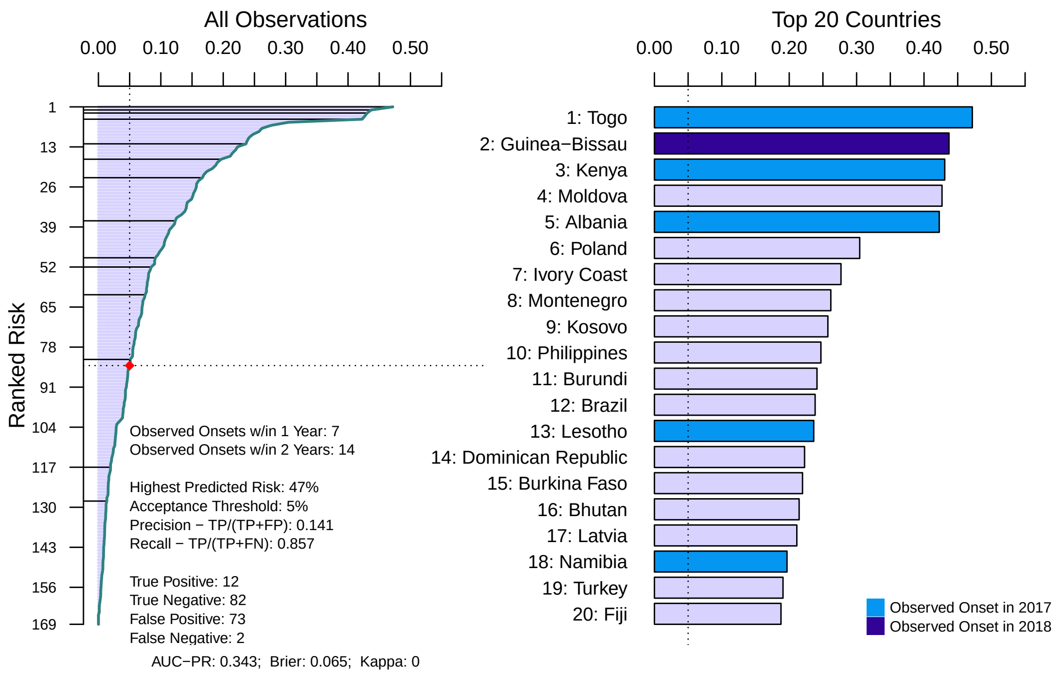


Figure 29: Random forest: 2019-2020 Forecast

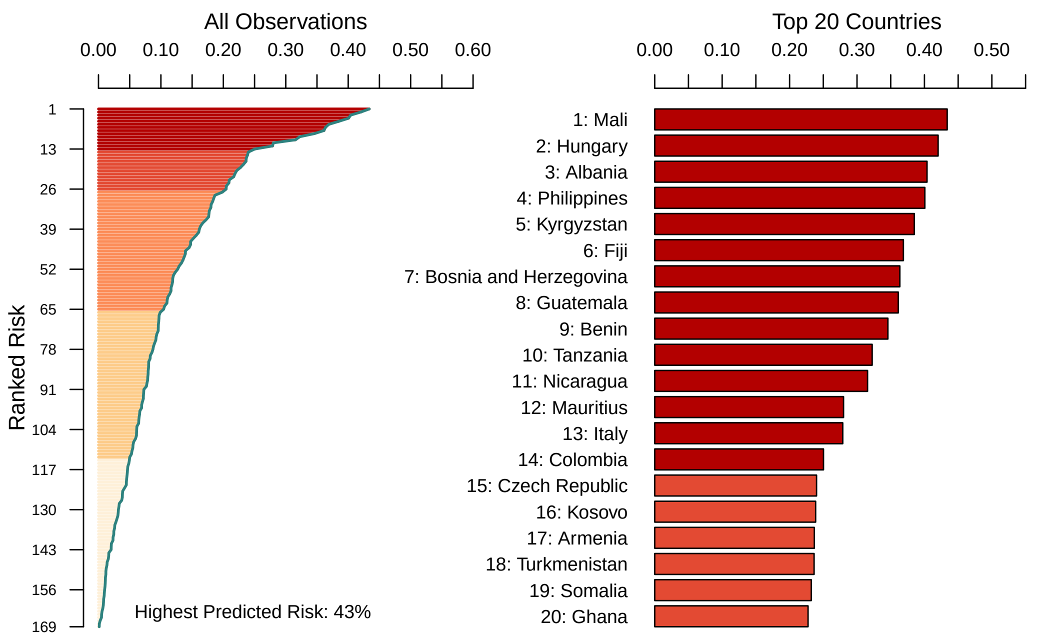


Figure 30: Gradient boosted forest: 2011-2012 Test Forecast

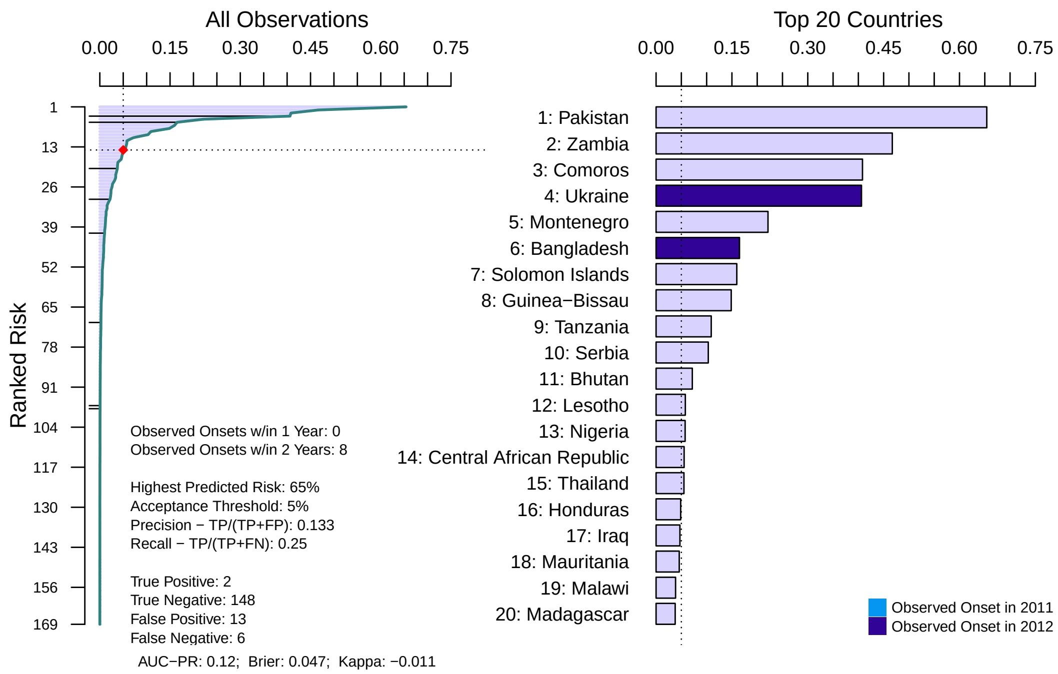


Figure 31: Gradient boosted forest: 2012-2013 Test Forecast

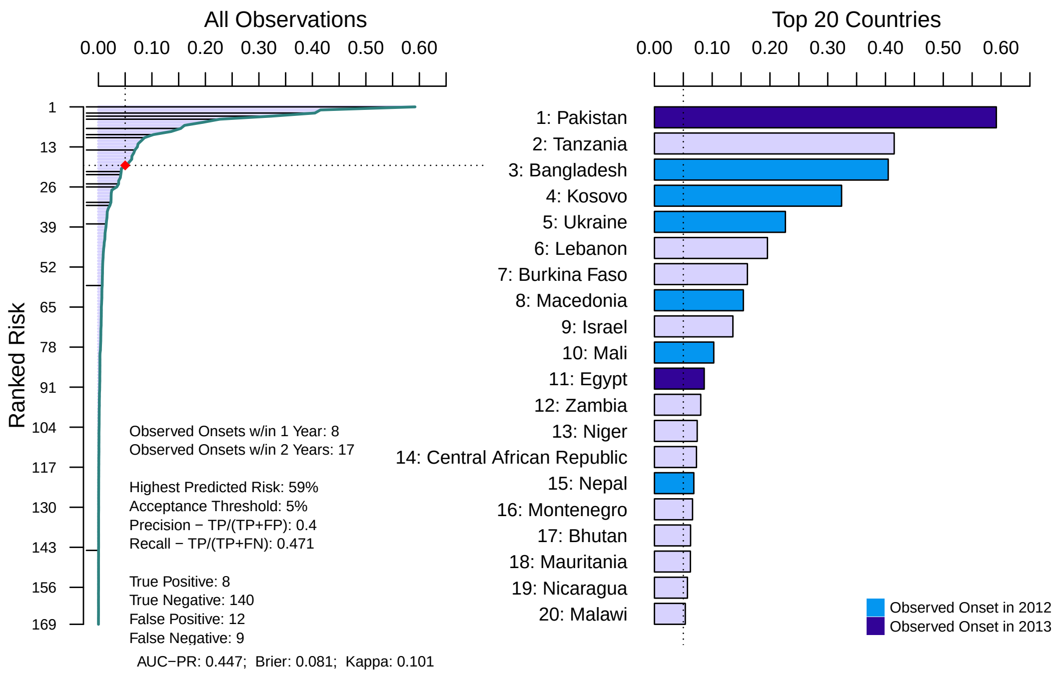


Figure 32: Gradient boosted forest: 2013-2014 Test Forecast

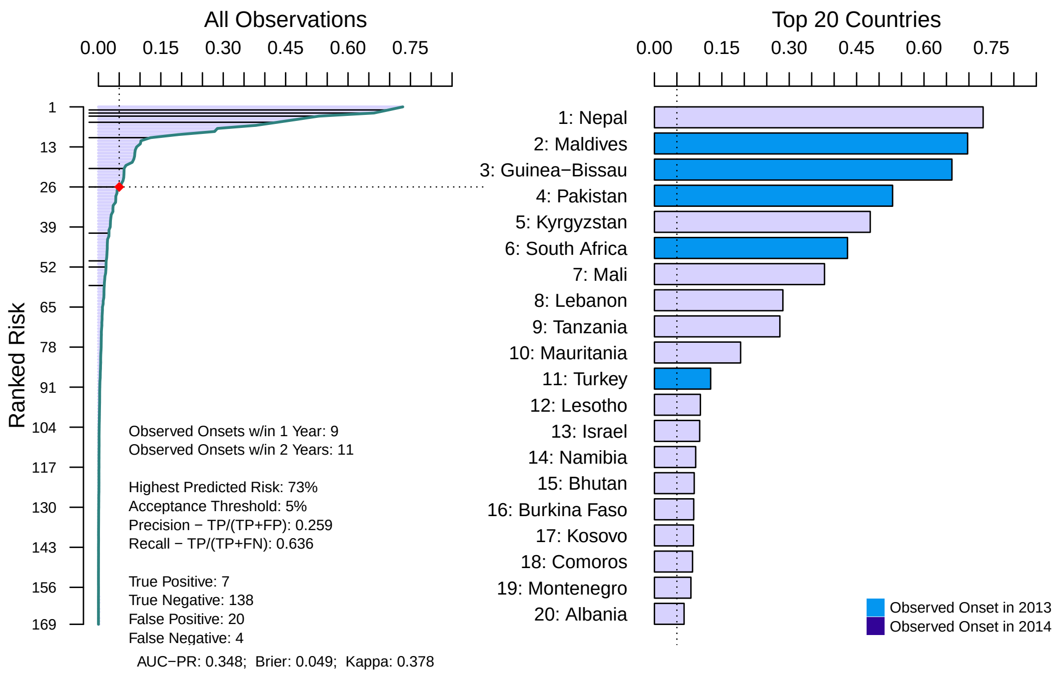


Figure 33: Gradient boosted forest: 2014-2015 Test Forecast

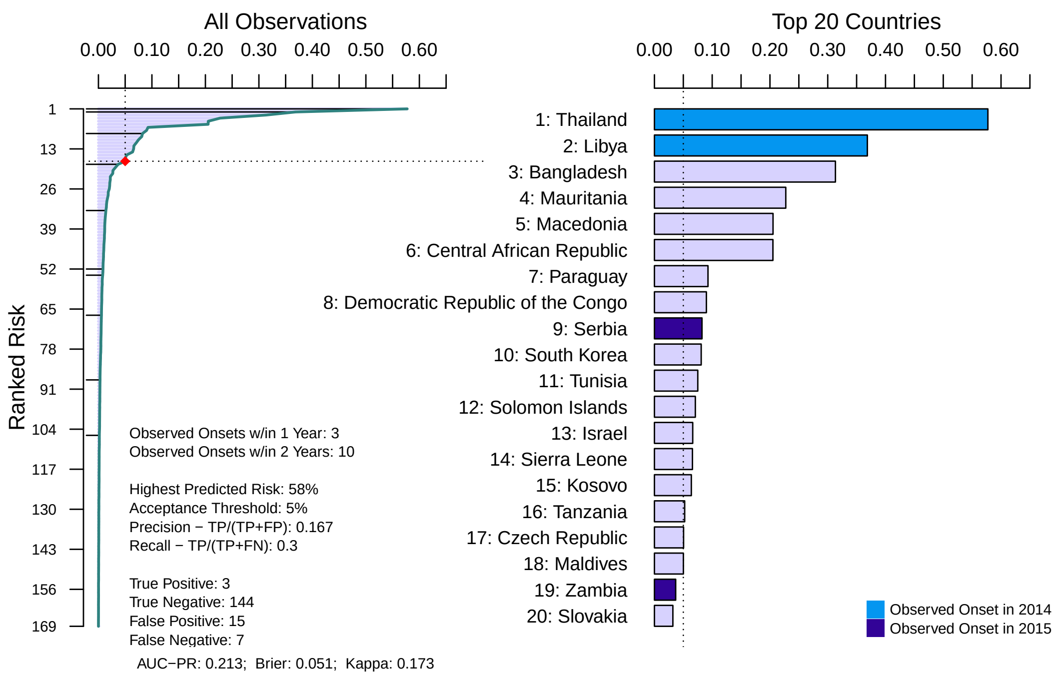


Figure 34: Gradient boosted forest: 2015-2016 Test Forecast

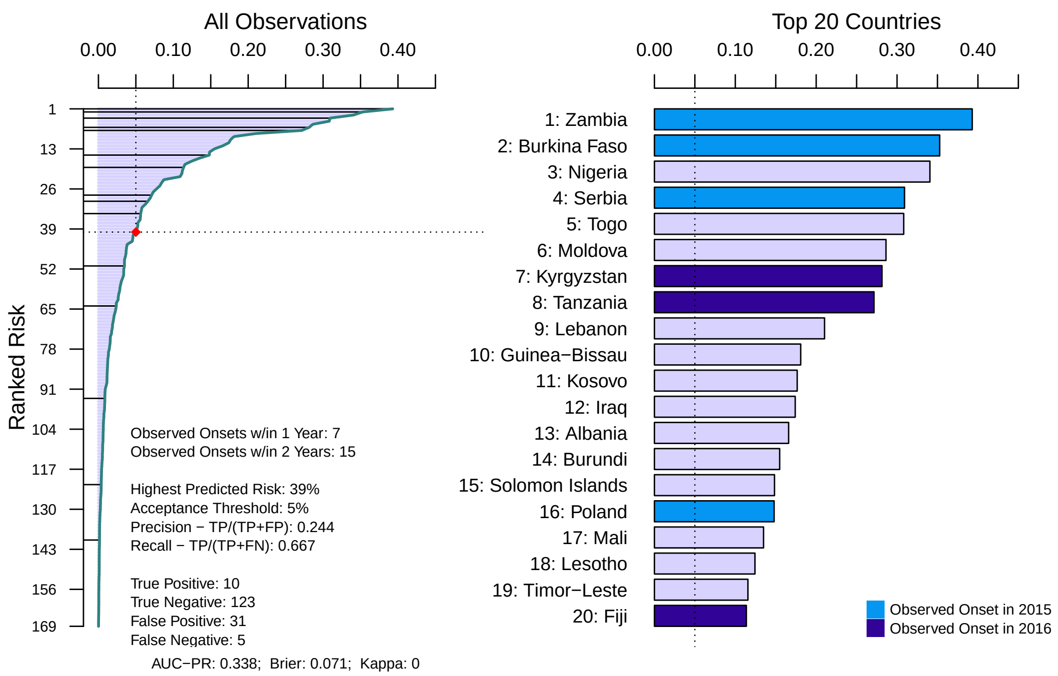


Figure 35: Gradient boosted forest: 2016-2017 Test Forecast

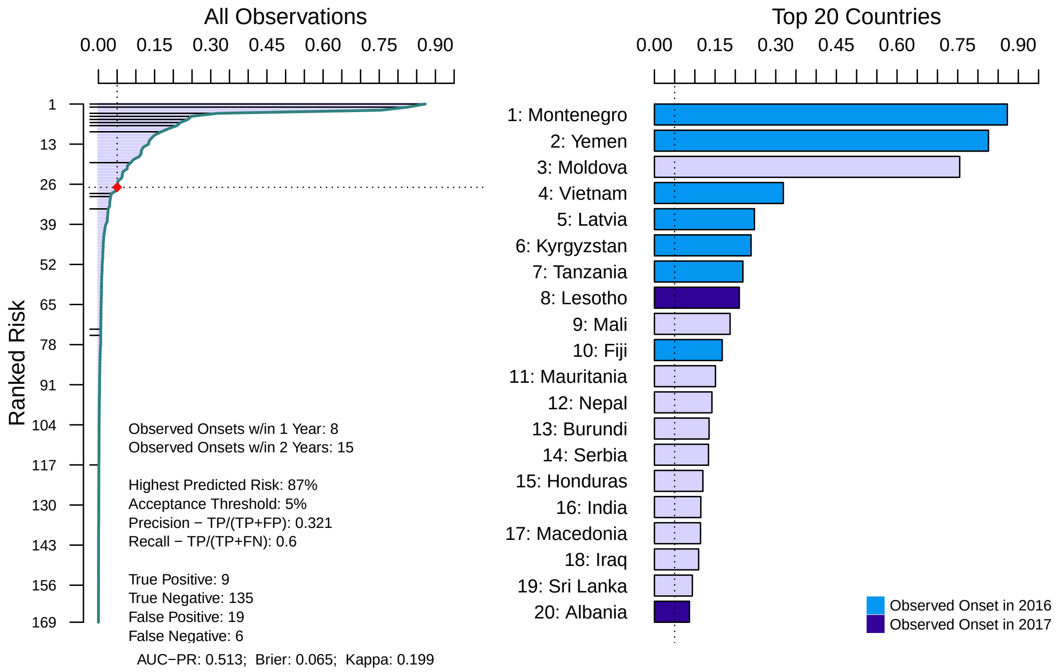


Figure 36: Gradient boosted forest: 2017-2018 Test Forecast

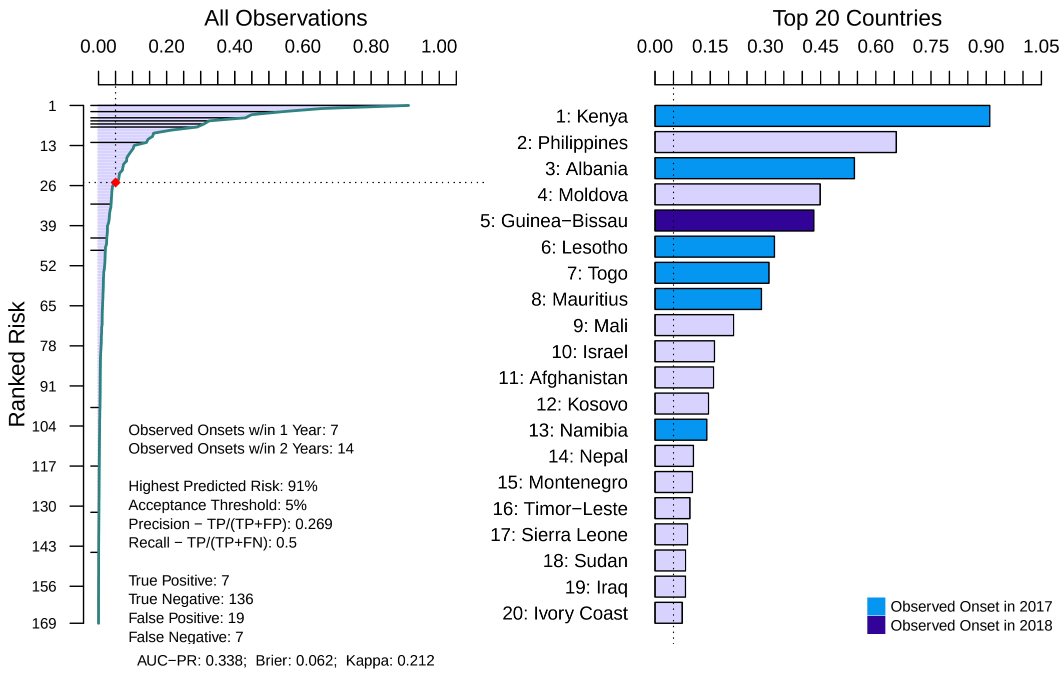
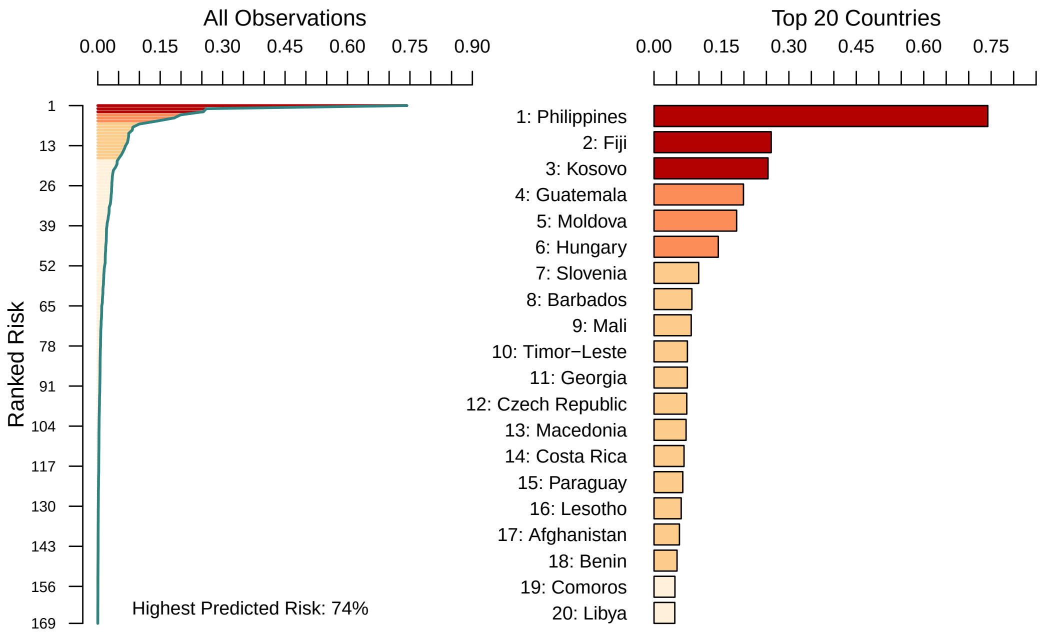


Figure 37: Gradient boosted forest: 2019-2020 Forecast



D Appendix: Adverse Regime Transitions

The tables below lists the 189 adverse regime transitions that are captured in our data by year. They provide the RoW classification before and after the ART event as well as the year that the event manifested.

Table 8: Adverse Regime Transitions – 1970-2018

	Country Name	RoW Classification before ART	RoW Classification after ART	Year of ART
1	Bolivia	Electoral Autocracy	Closed Autocracy	1970
2	Malaysia	Electoral Autocracy	Closed Autocracy	1970
3	Somalia	Electoral Autocracy	Closed Autocracy	1970
4	Cambodia	Electoral Autocracy	Closed Autocracy	1971
5	Lesotho	Electoral Autocracy	Closed Autocracy	1971
6	Ecuador	Electoral Autocracy	Closed Autocracy	1972
7	Madagascar	Electoral Autocracy	Closed Autocracy	1972
8	Philippines	Electoral Autocracy	Closed Autocracy	1972
9	Uruguay	Electoral Democracy	Electoral Autocracy	1972
10	Cambodia	Electoral Autocracy	Closed Autocracy	1973
11	Chile	Electoral Democracy	Electoral Autocracy	1973
12	Ghana	Electoral Autocracy	Closed Autocracy	1973
13	Honduras	Electoral Autocracy	Closed Autocracy	1973
14	Uruguay	Electoral Autocracy	Closed Autocracy	1973
15	Afghanistan	Electoral Autocracy	Closed Autocracy	1974
16	Chile	Electoral Autocracy	Closed Autocracy	1974
17	Jamaica	Electoral Democracy	Electoral Autocracy	1974
18	Nicaragua	Electoral Autocracy	Closed Autocracy	1974
19	Rwanda	Electoral Autocracy	Closed Autocracy	1974
20	Tunisia	Electoral Autocracy	Closed Autocracy	1974
21	Bangladesh	Electoral Autocracy	Closed Autocracy	1975
22	India	Electoral Democracy	Electoral Autocracy	1975
23	Laos	Electoral Autocracy	Closed Autocracy	1975
24	Argentina	Electoral Democracy	Electoral Autocracy	1976
25	Cape Verde	Electoral Autocracy	Closed Autocracy	1976
26	Israel	Liberal Democracy	Electoral Democracy	1976

Table 8: Adverse Regime Transitions – 1970-2018 (cont.)

	Country Name	RoW Classification before ART	RoW Classification after ART	Year of ART
27	Argentina	Electoral Autocracy	Closed Autocracy	1977
28	Colombia	Electoral Democracy	Electoral Autocracy	1977
29	Sudan	Electoral Autocracy	Closed Autocracy	1977
30	Thailand	Electoral Autocracy	Closed Autocracy	1977
31	Vietnam	Electoral Autocracy	Closed Autocracy	1977
32	Comoros	Electoral Autocracy	Closed Autocracy	1978
33	Malawi	Electoral Autocracy	Closed Autocracy	1978
34	Pakistan	Electoral Autocracy	Closed Autocracy	1978
35	The Gambia	Electoral Democracy	Electoral Autocracy	1978
36	Zimbabwe	Electoral Autocracy	Closed Autocracy	1978
37	Equatorial Guinea	Electoral Autocracy	Closed Autocracy	1979
38	El Salvador	Electoral Autocracy	Closed Autocracy	1980
39	Suriname	Electoral Democracy	Closed Autocracy	1980
40	Turkey	Electoral Democracy	Closed Autocracy	1980
41	Bolivia	Electoral Autocracy	Closed Autocracy	1981
42	Burkina Faso	Electoral Autocracy	Closed Autocracy	1981
43	Ghana	Electoral Democracy	Closed Autocracy	1981
44	Jamaica	Electoral Democracy	Electoral Autocracy	1981
45	Philippines	Electoral Autocracy	Closed Autocracy	1981
46	Solomon Islands	Electoral Democracy	Electoral Autocracy	1981
47	Djibouti	Electoral Autocracy	Closed Autocracy	1982
48	Senegal	Electoral Democracy	Electoral Autocracy	1982
49	Bangladesh	Electoral Autocracy	Closed Autocracy	1983
50	Guatemala	Electoral Autocracy	Closed Autocracy	1983
51	Colombia	Electoral Democracy	Electoral Autocracy	1984
52	Nigeria	Electoral Autocracy	Closed Autocracy	1984
53	Somalia	Electoral Autocracy	Closed Autocracy	1984
54	South Africa	Electoral Autocracy	Closed Autocracy	1985
55	Sudan	Electoral Autocracy	Closed Autocracy	1985
56	Uganda	Electoral Autocracy	Closed Autocracy	1985
57	Fiji	Electoral Democracy	Closed Autocracy	1987
58	Malawi	Electoral Autocracy	Closed Autocracy	1987
59	Burundi	Electoral Autocracy	Closed Autocracy	1988
60	Haiti	Electoral Autocracy	Closed Autocracy	1989

Table 8: Adverse Regime Transitions – 1970-2018 (cont.)

	Country Name	RoW Classification before ART	RoW Classification after ART	Year of ART
61	Dominican Republic	Electoral Democracy	Electoral Autocracy	1990
62	Liberia	Electoral Autocracy	Closed Autocracy	1990
63	Solomon Islands	Electoral Democracy	Electoral Autocracy	1990
64	Sudan	Electoral Autocracy	Closed Autocracy	1990
65	The Gambia	Electoral Democracy	Electoral Autocracy	1990
66	Laos	Electoral Autocracy	Closed Autocracy	1991
67	Thailand	Electoral Autocracy	Closed Autocracy	1991
68	Estonia	Electoral Democracy	Electoral Autocracy	1992
69	Haiti	Electoral Autocracy	Closed Autocracy	1992
70	Papua New Guinea	Electoral Democracy	Electoral Autocracy	1992
71	Peru	Electoral Democracy	Closed Autocracy	1992
72	Angola	Electoral Autocracy	Closed Autocracy	1993
73	Sierra Leone	Electoral Autocracy	Closed Autocracy	1993
74	Nigeria	Electoral Autocracy	Closed Autocracy	1994
75	Uganda	Electoral Autocracy	Closed Autocracy	1994
76	Armenia	Electoral Democracy	Electoral Autocracy	1995
77	Lesotho	Electoral Autocracy	Closed Autocracy	1995
78	The Gambia	Electoral Autocracy	Closed Autocracy	1995
79	Burundi	Electoral Autocracy	Closed Autocracy	1996
80	Niger	Electoral Democracy	Electoral Autocracy	1996
81	Russia	Electoral Democracy	Electoral Autocracy	1996
82	Belarus	Electoral Democracy	Electoral Autocracy	1997
83	El Salvador	Electoral Democracy	Electoral Autocracy	1997
84	Republic of the Congo	Electoral Autocracy	Closed Autocracy	1997
85	Zambia	Electoral Democracy	Electoral Autocracy	1997
86	Sierra Leone	Electoral Autocracy	Closed Autocracy	1998
87	Ukraine	Electoral Democracy	Electoral Autocracy	1998
88	Burkina Faso	Electoral Democracy	Electoral Autocracy	1999
89	Croatia	Electoral Democracy	Electoral Autocracy	1999
90	Lesotho	Electoral Autocracy	Closed Autocracy	1999
91	Pakistan	Electoral Autocracy	Closed Autocracy	1999
92	Comoros	Electoral Autocracy	Closed Autocracy	2000
93	Fiji	Electoral Democracy	Closed Autocracy	2000
94	Iraq	Electoral Autocracy	Closed Autocracy	2000
95	Macedonia	Electoral Democracy	Electoral Autocracy	2000
96	Solomon Islands	Electoral Democracy	Electoral Autocracy	2000
97	Thailand	Electoral Democracy	Electoral Autocracy	2000

Table 8: Adverse Regime Transitions – 1970-2018 (cont.)

	Country Name	RoW Classification before ART	RoW Classification after ART	Year of ART
98	Madagascar	Electoral Democracy	Electoral Autocracy	2001
99	Tanzania	Electoral Democracy	Electoral Autocracy	2001
100	Nepal	Electoral Autocracy	Closed Autocracy	2002
101	Burkina Faso	Electoral Democracy	Electoral Autocracy	2003
102	Nigeria	Electoral Democracy	Electoral Autocracy	2003
103	Central African Republic	Electoral Autocracy	Closed Autocracy	2004
104	Liberia	Electoral Autocracy	Closed Autocracy	2004
105	Malawi	Electoral Democracy	Electoral Autocracy	2004
106	Philippines	Electoral Democracy	Electoral Autocracy	2004
107	Haiti	Electoral Autocracy	Closed Autocracy	2005
108	Kosovo	Electoral Democracy	Electoral Autocracy	2005
109	Moldova	Electoral Democracy	Electoral Autocracy	2005
110	Sri Lanka	Electoral Democracy	Electoral Autocracy	2005
111	Zambia	Electoral Democracy	Electoral Autocracy	2005
112	Bangladesh	Electoral Democracy	Electoral Autocracy	2006
113	Hungary	Liberal Democracy	Electoral Democracy	2006
114	Mauritania	Electoral Autocracy	Closed Autocracy	2006
115	Montenegro	Electoral Democracy	Electoral Autocracy	2006
116	Solomon Islands	Electoral Democracy	Electoral Autocracy	2006
117	Thailand	Electoral Democracy	Closed Autocracy	2006
118	Venezuela	Electoral Democracy	Electoral Autocracy	2006
119	Bangladesh	Electoral Autocracy	Closed Autocracy	2007
120	Fiji	Electoral Democracy	Closed Autocracy	2007
121	Nicaragua	Electoral Democracy	Electoral Autocracy	2007
122	South Africa	Liberal Democracy	Electoral Democracy	2007
123	Guinea-Bissau	Electoral Democracy	Electoral Autocracy	2008
124	Madagascar	Electoral Democracy	Electoral Autocracy	2008
125	Mauritania	Electoral Autocracy	Closed Autocracy	2008
126	Moldova	Electoral Democracy	Electoral Autocracy	2008
127	Guinea	Electoral Autocracy	Closed Autocracy	2009
128	Honduras	Electoral Democracy	Electoral Autocracy	2009
129	Niger	Electoral Democracy	Electoral Autocracy	2009
130	Tanzania	Electoral Democracy	Electoral Autocracy	2009

Table 8: Adverse Regime Transitions – 1970-2018 (cont.)

	Country Name	RoW Classification before ART	RoW Classification after ART	Year of ART
131	Guinea-Bissau	Electoral Democracy	Electoral Autocracy	2010
132	Hungary	Liberal Democracy	Electoral Democracy	2010
133	Iraq	Electoral Democracy	Electoral Autocracy	2010
134	Israel	Liberal Democracy	Electoral Democracy	2010
135	Lebanon	Electoral Democracy	Electoral Autocracy	2010
136	Madagascar	Electoral Autocracy	Closed Autocracy	2010
137	Montenegro	Electoral Democracy	Electoral Autocracy	2010
138	Niger	Electoral Autocracy	Closed Autocracy	2010
139	Taiwan	Liberal Democracy	Electoral Democracy	2010
140	Togo	Electoral Democracy	Electoral Autocracy	2010
141	Bangladesh	Electoral Democracy	Electoral Autocracy	2012
142	Kosovo	Electoral Democracy	Electoral Autocracy	2012
143	Macedonia	Electoral Democracy	Electoral Autocracy	2012
144	Mali	Electoral Democracy	Electoral Autocracy	2012
145	Namibia	Liberal Democracy	Electoral Democracy	2012
146	Nepal	Electoral Democracy	Electoral Autocracy	2012
147	Turkmenistan	Electoral Autocracy	Closed Autocracy	2012
148	Ukraine	Electoral Democracy	Electoral Autocracy	2012
149	Egypt	Electoral Autocracy	Closed Autocracy	2013
150	Guinea-Bissau	Electoral Autocracy	Closed Autocracy	2013
151	Maldives	Electoral Democracy	Electoral Autocracy	2013
152	Pakistan	Electoral Democracy	Electoral Autocracy	2013
153	Slovakia	Liberal Democracy	Electoral Democracy	2013
154	South Africa	Liberal Democracy	Electoral Democracy	2013
155	Syria	Electoral Autocracy	Closed Autocracy	2013
156	Thailand	Electoral Democracy	Electoral Autocracy	2013
157	Turkey	Electoral Democracy	Electoral Autocracy	2013
158	Ivory Coast	Electoral Democracy	Electoral Autocracy	2014
159	Libya	Electoral Democracy	Closed Autocracy	2014
160	Thailand	Electoral Autocracy	Closed Autocracy	2014
161	Benin	Liberal Democracy	Electoral Democracy	2015
162	Burkina Faso	Electoral Democracy	Electoral Autocracy	2015
163	Comoros	Electoral Democracy	Electoral Autocracy	2015
164	Ghana	Liberal Democracy	Electoral Democracy	2015
165	Poland	Liberal Democracy	Electoral Democracy	2015
166	Serbia	Electoral Democracy	Electoral Autocracy	2015
167	Zambia	Electoral Democracy	Electoral Autocracy	2015

Table 8: Adverse Regime Transitions – 1970-2018 (cont.)

	Country Name	RoW Classification before ART	RoW Classification after ART	Year of ART
168	Fiji	Electoral Democracy	Electoral Autocracy	2016
169	Kyrgyzstan	Electoral Democracy	Electoral Autocracy	2016
170	Latvia	Liberal Democracy	Electoral Democracy	2016
171	Lithuania	Liberal Democracy	Electoral Democracy	2016
172	Montenegro	Electoral Democracy	Electoral Autocracy	2016
173	Tanzania	Electoral Democracy	Electoral Autocracy	2016
174	Vietnam	Electoral Autocracy	Closed Autocracy	2016
175	Yemen	Electoral Autocracy	Closed Autocracy	2016
176	Albania	Liberal Democracy	Electoral Democracy	2017
177	Botswana	Liberal Democracy	Electoral Democracy	2017
178	Kenya	Electoral Democracy	Electoral Autocracy	2017
179	Lesotho	Electoral Democracy	Electoral Autocracy	2017
180	Mauritius	Liberal Democracy	Electoral Democracy	2017
181	Namibia	Liberal Democracy	Electoral Democracy	2017
182	Togo	Electoral Democracy	Electoral Autocracy	2017
183	Cape Verde	Liberal Democracy	Electoral Democracy	2018
184	Chile	Liberal Democracy	Electoral Democracy	2018
185	Greece	Liberal Democracy	Electoral Democracy	2018
186	Guinea-Bissau	Electoral Democracy	Electoral Autocracy	2018
187	Lebanon	Electoral Democracy	Electoral Autocracy	2018
188	South Korea	Liberal Democracy	Electoral Democracy	2018
189	Tunisia	Liberal Democracy	Electoral Democracy	2018

E Appendix: Data Description

The table below provides a list of all of the variables we use in our various models. Most variable names are simply the variable tag from the V-Dem data set. Please see the V-Dem code book for a full description of these variables. The other variable names are self explanatory. This table also provides basic descriptive statistics: the minimum and maximum value, the mean, and the standard deviation of each variable in our sample.

Table 9: Variables and Descriptive Statistics

Variable	Min	Max	Mean	SD
is_closed_autocracy	0	1	0.306	0.461
currentRegimeDuration	1	118	26.417	28.042
low_border_case	0	1	0.139	0.346
high_border_case	0	1	0.113	0.316
yrs_since_any_neg_change	0	117	40.521	34.302
num_of_neg_changes	0	10	1.369	1.642
any_neg_change	0	1	0.024	0.154
num_of_neg_changes_3yrs	0	2	0.072	0.266
num_of_neg_changes_5yrs	0	2	0.120	0.345
num_of_neg_changes_10yrs	0	3	0.232	0.496
was_neg_change_last_3yrs	0	1	0.070	0.254
was_neg_change_last_5yrs	0	1	0.113	0.316
was_neg_change_last_10yrs	0	1	0.200	0.400
v2x_regime.0	0	1	0.306	0.461
v2x_regime.1	0	1	0.278	0.448
v2x_regime.2	0	1	0.208	0.406
v2x_regime.3	0	1	0.208	0.406
v2x_regime_amb.0	0	1	0.267	0.442
v2x_regime_amb.1	0	1	0.040	0.195
v2x_regime_amb.2	0	1	0.029	0.169
v2x_regime_amb.3	0	1	0.212	0.409
v2x_regime_amb.4	0	1	0.037	0.188
v2x_regime_amb.5	0	1	0.049	0.217
v2x_regime_amb.6	0	1	0.123	0.328
v2x_regime_amb.7	0	1	0.036	0.187
v2x_regime_amb.8	0	1	0.060	0.238
v2x_regime_amb.9	0	1	0.147	0.354

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
is_leg	0	1	0.912	0.284
is_elec	0	1	0.854	0.353
is_election_year	0	1	0.261	0.439
v2elrgstry	-3.967	2.872	0.286	1.313
v2elvotbuy	-2.989	3.400	0.033	1.351
v2elirreg	-3.081	3.109	0.041	1.369
v2elintim	-4.037	3.445	-0.055	1.458
v2elpeace	-4.578	2.550	0.121	1.302
v2elfrfair	-3.410	2.883	-0.037	1.527
v2elmulpar	-3.534	2.538	0.080	1.476
v2elboycot	-4.317	1.927	0.078	1.228
v2elaccept	-3.805	2.687	-0.062	1.362
v2elasmoff	-6.669	0.885	-0.076	1.066
v2eldonate	-2.561	3.895	-0.227	1.332
v2elpubfin	-2.586	3.770	-0.025	1.451
v2ellocons	0	59	8.052	8.907
v2ellocumul	0	59	14.617	10.428
v2elprescons	0	30	2.673	4.326
v2elprescumul	0	36	4.807	6.696
v2elpaidig	-2.580	3.161	0.056	1.328
v2elfrcamp	-2.241	2.946	0.346	1.360
v2elpdcamp	-2.727	3.385	0.058	1.391
v2elmonref	0	1	0.029	0.168
v2elmonden	0	1	0.025	0.156
v2x_polyarchy	0.010	0.948	0.447	0.290
v2x_freexp_altinf	0.012	0.990	0.569	0.331
v2x_frassoc_thick	0.019	0.957	0.547	0.340
v2xel_frefair	0	0.993	0.469	0.347
v2x_elecoff	0	1	0.762	0.411
v2x_liberal	0.004	0.985	0.534	0.292
v2xcl_rol	0	0.994	0.602	0.300
v2x_jucon	0.003	0.991	0.534	0.310
v2xlg_legcon	0	0.984	0.470	0.345
v2x_partip	0.018	0.887	0.408	0.207
v2x_cspart	0.021	0.989	0.575	0.283
v2xdl_delib	0.009	0.988	0.555	0.300
v2x_egal	0.054	0.983	0.588	0.227
v2xeg_eqprotec	0.011	0.992	0.611	0.258
v2xeg_eqaccess	0.025	0.991	0.580	0.246

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
v2xeg_eqdr	0.010	0.985	0.578	0.283
v2elembaut	-2.642	3.818	0.311	1.706
v2elembcap	-3.002	3.654	0.415	1.488
v2elreggov	0	1	0.778	0.416
v2ellocgov	0	1	0.958	0.200
v2elrsthos	0	1	0.879	0.326
v2elrstrct	0	1	0.905	0.293
v2psparban	-3.700	2.891	0.265	1.792
v2psbars	-3.735	3.091	0.476	1.708
v2psoppaut	-3.542	3.534	0.369	1.944
v2psorgs	-3.097	3.213	0.671	1.405
v2psprbrch	-3.191	3.545	0.658	1.369
v2psprlnks	-3.205	3.509	0.113	1.382
v2psplats	-3.163	3.349	0.379	1.635
v2pscnslnl	-2.685	4.606	0.140	1.357
v2pscohesv	-3.690	2.557	0.691	1.219
v2pscomprg	-3.467	2.582	0.457	1.199
v2psnatpar	-2.772	3.017	0.125	1.415
v2pssunpar	-2.751	2.798	-0.010	1.594
v2exremhsp	-2.495	4.164	-0.315	1.199
v2exdfdshs	-3.490	3.029	-0.178	1.502
v2exdfcbhs	-3.377	2.456	0.184	1.497
v2exdfvths	-3.396	2.525	-0.091	1.496
v2exdfdmhs	-3.281	2.402	0.159	1.637
v2exdfpphs	-2.553	3.328	0.327	1.477
v2exhoshog	0	1	0.399	0.490
v2exrescon	-3.223	3.476	0.293	1.429
v2exbribe	-3.142	3.607	0.012	1.537
v2exembez	-3.232	3.570	0.050	1.544
v2excrptps	-3.074	4.005	-0.150	1.505
v2exthftps	-3.142	3.712	0.049	1.518
v2ex_elechos	0	1	0.460	0.498
v2ex_hogw	0	1	0.381	0.471
v2expathhs	0	8	5.208	2.380
v2lgbicam	0	2	1.251	0.604
v2lgqstexp	-2.336	2.320	0.248	1.336
v2lginvstp	-2.939	3.844	0.217	1.481
v2lgotovst	-2.974	3.124	0.181	1.397
v2lgcrrpt	-3.284	3.322	-0.087	1.321
v2lgoppart	-2.528	3.507	0.285	1.582
v2lgfunds	-2.427	2.418	0.214	1.291
v2lgdsadlobin	-9.368	0.605	-0.063	0.975
v2lglegplo	-3.778	1.963	0.352	1.196
v2lgcomslo	-3.627	3.742	0.632	1.097
v2lgsrvlo	-2.213	2.154	-0.051	1.361

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
v2ex_hosw	0	1	0.619	0.471
v2lgamend	0	1	0.426	0.495
v2dlreason	-2.911	3.716	0.439	1.347
v2dlcommon	-3.568	2.858	0.409	1.188
v2dlcountr	-3.359	3.548	0.201	1.435
v2dlconslt	-3.051	4.455	0.555	1.414
v2dlengage	-3.123	3.457	0.460	1.426
v2dlencmps	-3.452	3.438	0.545	1.238
v2dlunivl	-3.289	3.407	0.661	1.142
v2jureform	-3.607	3.398	0.024	1.167
v2jupurge	-3.830	2.847	0.427	1.262
v2jupoatck	-4.449	3.048	0.334	1.230
v2jupack	-4.454	1.710	-0.056	1.242
v2juaccnt	-3.087	3.642	0.495	1.278
v2jucorrdc	-3.161	3.290	0.043	1.512
v2juhcind	-3.205	3.471	0.100	1.469
v2juncind	-3.375	3.359	0.259	1.470
v2juhccomp	-3.910	2.875	0.281	1.468
v2jucomp	-3.458	3.264	0.206	1.471
v2jureview	-2.719	1.842	0.180	1.241
v2clacfree	-3.506	3.798	0.617	1.669
v2clrelig	-3.929	2.798	0.632	1.468
v2cltort	-3.067	3.658	0.388	1.631
v2clkill	-3.511	3.514	0.638	1.667
v2cltrnslw	-3.730	4.170	0.476	1.534
v2clrspct	-3.685	4.455	0.202	1.512
v2clfmove	-4.206	2.949	0.645	1.474
v2cldmovem	-5.016	2.689	0.745	1.288
v2cldmovew	-4.756	3.238	0.696	1.402
v2cldiscm	-3.781	3.880	0.508	1.681
v2cldiscw	-3.534	3.500	0.493	1.582
v2clslavem	-4.100	3.033	0.883	1.130
v2clslavef	-4.281	2.973	0.682	1.119
v2clstown	-4.197	3.295	0.107	1.365
v2clprptym	-4.398	2.425	0.643	1.267
v2clprptyw	-3.750	2.822	0.609	1.337
v2clacjstm	-4.056	3.896	0.546	1.476
v2clacjstw	-3.974	3.722	0.492	1.466
v2clacjust	-2.897	3.594	0.836	1.197
v2clsocgrp	-2.996	3.102	0.526	1.305
v2clrgunev	-2.942	2.794	0.347	1.301
v2svdomaut	-3.254	2.169	1.036	0.730
v2svinlaut	-3.147	2.733	1.097	0.769
v2svstterr	33.750	100	91.271	10.554
v2cseeorgs	-3.226	3.548	0.498	1.592

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
v2csreprss	-3.729	3.379	0.504	1.628
v2cscnsult	-2.454	3.848	0.400	1.476
v2csprtcpt	-3.532	3.263	0.333	1.470
v2csgender	-3.499	3.238	0.788	1.087
v2csantimv	-2.973	4.015	-0.515	1.265
v2csrlgrep	-4.121	2.885	0.579	1.455
v2csrlgcon	-2.854	3.054	0.159	1.335
v2mecenefm	-3.089	3.569	0.220	1.671
v2mecrit	-3.314	3.595	0.347	1.698
v2merange	-3.111	3.171	0.285	1.629
v2meharjrn	-3.088	3.985	0.284	1.628
v2meslfcen	-3.241	3.268	0.211	1.553
v2mebias	-3.585	3.732	0.264	1.702
v2mecorrpt	-3.179	3.456	0.162	1.645
v2pepwrses	-2.950	2.988	0.478	1.141
v2pepwrsoc	-2.638	3.399	0.543	1.269
v2pepwrgen	-2.884	3.876	0.545	1.199
v2pepwrort	-2.204	3.485	-0.170	1.242
v2peedueq	-3.102	3.634	0.398	1.482
v2pehealth	-3.271	3.689	0.434	1.503
v2x_accountability	-1.981	2.175	0.352	1.020
v2x_veracc	-1.579	1.893	0.415	0.877
v2x_diagacc	-2.145	2.178	0.364	1.046
v2x_horacc	-2.144	2.385	0.194	1.048
v2xex_elecleg	0	1	0.840	0.335
v2x_civlib	0.009	0.978	0.596	0.288
v2x_clphy	0.013	0.987	0.580	0.312
v2x_clpol	0.009	0.987	0.580	0.334
v2x_clpriv	0	0.973	0.630	0.287
v2x_corr	0.006	0.976	0.513	0.296
v2x_EDcomp_thick	0.002	0.958	0.509	0.287
v2x_elecreg	0	1	0.854	0.353
v2x_freexp	0.011	0.991	0.570	0.324
v2x_gencl	0.001	0.985	0.612	0.273
v2x_gencs	0.014	0.975	0.603	0.246
v2x_hosabort	0	1	0.002	0.039
v2x_hosinter	0	1	0.008	0.090
v2x_pubcorr	0.004	0.979	0.491	0.300
v2x_rule	0.005	0.998	0.520	0.310
v2xcl_acjst	0.002	0.995	0.582	0.286
v2xcl_disc	0.007	0.992	0.584	0.315
v2xcl_dmove	0	0.970	0.634	0.262
v2xcl_prpty	0.002	0.949	0.600	0.269
v2xcl_slave	0.001	0.969	0.635	0.240
v2xcs_ccsi	0.008	0.977	0.576	0.314

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
v2xel_elecparl	0	1	0.216	0.412
v2xel_elecpres	0	1	0.103	0.305
v2xex_elecreg	0	1	0.488	0.500
v2xlg_elecreg	0	1	0.852	0.355
v2x_ex_confidence	0	1	0.329	0.378
v2x_ex_direlect	0	1	0.366	0.475
v2x_ex_hereditary	0	1	0.043	0.159
v2x_ex_military	0	1	0.185	0.236
v2x_ex_party	0	1	0.148	0.210
v2x_execorr	0.011	0.978	0.494	0.301
v2x_legabort	0	1	0.002	0.042
v2xlg_leginter	0	1	0.020	0.140
v2x_neopat	0.006	0.990	0.508	0.308
v2xnp_client	0.012	0.986	0.485	0.259
v2xnp_pres	0.010	0.989	0.471	0.322
v2xnp_regcorr	0.006	0.980	0.499	0.304
diff_year_prior_v2elrgstry	-2.725	3.553	0.015	0.232
diff_year_prior_v2elvotbuy	-2.877	2.771	-0.002	0.241
diff_year_prior_v2elirreg	-3.293	3.108	0.005	0.274
diff_year_prior_v2elintim	-3.299	4.569	0.013	0.296
diff_year_prior_v2elpeace	-3.617	3.779	0.004	0.295
diff_year_prior_v2elfrfair	-3.669	4.702	0.014	0.337
diff_year_prior_v2elmulpar	-4.258	4.612	0.022	0.345
diff_year_prior_v2elboycot	-4.780	4.759	0.004	0.379
diff_year_prior_v2elaccept	-4.225	4.246	0.009	0.328
diff_year_prior_v2elasmoff	-7.222	6.927	0.005	0.398
diff_year_prior_v2eldonate	-2.113	3.169	0.034	0.215
diff_year_prior_v2elpubfin	-1.543	3.918	0.024	0.215
diff_year_prior_v2ellocons	-25	3	0.149	0.933
diff_year_prior_v2ellocumul	0	3	0.231	0.463
diff_year_prior_v2elprescons	-19	3	0.095	0.576
diff_year_prior_v2elprescumul	0	3	0.126	0.395
diff_year_prior_v2elpaidig	-3.108	4.137	0.018	0.258
diff_year_prior_v2elfrcamp	-3.996	4.508	0.023	0.294
diff_year_prior_v2elpdcamp	-2.968	3.695	0.021	0.267
diff_year_prior_v2elmonref	-1	1	0.001	0.104
diff_year_prior_v2elmonden	-1	1	0.001	0.092
diff_year_prior_v2x_polyarchy	-0.475	0.702	0.005	0.046
diff_year_prior_v2x_freexp_altinf	-0.571	0.848	0.005	0.057
diff_year_prior_v2x_frassoc_thick	-0.669	0.827	0.006	0.060
diff_year_prior_v2xel_frefair	-0.865	0.919	0.005	0.083
diff_year_prior_v2x_elecoff	-1	1	0.006	0.169
diff_year_prior_v2x_liberal	-0.474	0.653	0.004	0.047

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
diff_year_prior_v2xcl_rol	-0.532	0.750	0.004	0.046
diff_year_prior_v2x_jucon	-0.622	0.822	0.002	0.050
diff_year_prior_v2xlg_legcon	-0.868	0.939	0.005	0.092
diff_year_prior_v2x_partip	-0.356	0.525	0.004	0.033
diff_year_prior_v2x_cspart	-0.508	0.711	0.006	0.052
diff_year_prior_v2xdl_delib	-0.667	0.821	0.005	0.061
diff_year_prior_v2x_egal	-0.240	0.438	0.002	0.032
diff_year_prior_v2xeg_eqprotec	-0.545	0.662	0.002	0.038
diff_year_prior_v2xeg_eqaccess	-0.340	0.553	0.003	0.040
diff_year_prior_v2xeg_eqdr	-0.354	0.520	0.001	0.031
diff_year_prior_v2elembaut	-3.086	3.922	0.025	0.272
diff_year_prior_v2elembcap	-2.259	3.904	0.017	0.216
diff_year_prior_v2elreggov	-1	1	0.001	0.045
diff_year_prior_v2ellocgov	-1	1	0.001	0.038
diff_year_prior_v2elrsthos	-1	1	-0.002	0.083
diff_year_prior_v2elrstrct	-1	1	0.001	0.062
diff_year_prior_v2psparban	-4.146	4.752	0.031	0.361
diff_year_prior_v2psbars	-3.495	4.394	0.027	0.319
diff_year_prior_v2psoppaut	-3.628	4.485	0.027	0.345
diff_year_prior_v2psorgs	-3.490	3.654	0.013	0.236
diff_year_prior_v2psprbrch	-3.340	3.156	0.008	0.222
diff_year_prior_v2psprlnks	-2.672	2.899	0.010	0.209
diff_year_prior_v2psplats	-2.524	3.900	0.014	0.222
diff_year_prior_v2pscnslnl	-3.354	3.022	0.018	0.216
diff_year_prior_v2pscohesv	-1.992	2.384	0	0.195
diff_year_prior_v2pscomprg	-3.269	3.627	0.005	0.224
diff_year_prior_v2psnatpar	-3.754	2.976	-0.003	0.357
diff_year_prior_v2pssunpar	-3.920	4.109	0.019	0.272
diff_year_prior_v2exremhsp	-3.736	3.063	0.006	0.279
diff_year_prior_v2exdfdshs	-4.133	2.845	-0.009	0.265
diff_year_prior_v2exdfcbhs	-3.227	3.322	-0.003	0.231
diff_year_prior_v2exdfvths	-3.028	3.279	-0.007	0.236
diff_year_prior_v2exdfdmhs	-3.283	3.039	-0.005	0.238
diff_year_prior_v2exdfpphs	-2.655	4.069	0.010	0.225
diff_year_prior_v2exhoshog	-1	1	-0.002	0.148
diff_year_prior_v2exrescon	-3.962	5.174	0.007	0.291
diff_year_prior_v2exbribe	-3.809	3.285	-0.001	0.237
diff_year_prior_v2exembez	-2.847	3.119	-0.001	0.243
diff_year_prior_v2excrptps	-2.522	2.919	-0.006	0.191
diff_year_prior_v2exthftps	-2.650	2.803	-0.005	0.211
diff_year_prior_v2ex_elechos	-1	1	0.009	0.168
diff_year_prior_v2ex_hogw	-1	1	0.001	0.105
diff_year_prior_v2expathhs	-8	8	0.038	1.002
diff_year_prior_v2lgbicam	-2	2	0.006	0.272
diff_year_prior_v2lgqstexp	-2.148	4.149	0.014	0.329

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
diff_year_prior_v2lginvstp	-2.661	3.769	0.015	0.337
diff_year_prior_v2lgotovst	-2.886	3.431	0.015	0.322
diff_year_prior_v2lgcrrpt	-3.966	3.941	-0.008	0.305
diff_year_prior_v2lgoppart	-2.730	4.508	0.017	0.366
diff_year_prior_v2lgfunds	-3.582	3.419	0.015	0.303
diff_year_prior_v2lgdsadlobin	-9.368	9.852	0.001	0.330
diff_year_prior_v2lglegplo	-3.600	4.463	0.011	0.310
diff_year_prior_v2lgcomslo	-3.627	4.796	0.021	0.288
diff_year_prior_v2lgsrvlo	-3.173	2.313	-0.006	0.277
diff_year_prior_v2ex_hosw	-1	1	-0.001	0.105
diff_year_prior_v2lgamend	-1	1	0.002	0.134
diff_year_prior_v2dlreason	-3.349	3.548	0.020	0.285
diff_year_prior_v2dlcommon	-2.238	3.033	0.013	0.248
diff_year_prior_v2dlcountr	-3.128	5.167	0.019	0.320
diff_year_prior_v2dlconslt	-3.315	4.817	0.020	0.312
diff_year_prior_v2dlengage	-3.446	3.811	0.023	0.306
diff_year_prior_v2dlencmps	-2.989	2.427	0.009	0.240
diff_year_prior_v2dlunivl	-3.384	4.001	0.007	0.215
diff_year_prior_v2jureform	-4.809	4.615	0.005	0.518
diff_year_prior_v2jupurge	-3.846	2.964	0.012	0.338
diff_year_prior_v2jupoatck	-4.557	3.449	-0.003	0.321
diff_year_prior_v2jupack	-3.524	3.724	0.007	0.311
diff_year_prior_v2juaccnt	-3.485	2.514	0.009	0.212
diff_year_prior_v2jucorrdc	-2.798	2.921	-0.005	0.184
diff_year_prior_v2juhcind	-2.791	4.524	0.012	0.265
diff_year_prior_v2juncind	-2.590	4.365	0.009	0.238
diff_year_prior_v2juhccomp	-2.630	3.685	0.006	0.225
diff_year_prior_v2jucomp	-3.511	3.540	0.006	0.224
diff_year_prior_v2jureview	-2.607	3.391	0.021	0.238
diff_year_prior_v2clacfree	-4.323	4.146	0.021	0.304
diff_year_prior_v2clrelig	-3.592	3.765	0.014	0.234
diff_year_prior_v2cltort	-3.661	3.753	0.021	0.295
diff_year_prior_v2clkill	-3.669	4.046	0.021	0.308
diff_year_prior_v2cltrnslw	-3.800	3.261	0.013	0.262
diff_year_prior_v2clrspct	-3.727	4.213	0.007	0.252
diff_year_prior_v2clfmove	-3.426	4.747	0.021	0.259
diff_year_prior_v2cldmovem	-4.805	5.435	0.015	0.247
diff_year_prior_v2cldmovew	-4.651	4.495	0.017	0.222
diff_year_prior_v2cldiscm	-3.696	4.530	0.025	0.325
diff_year_prior_v2cldiscw	-3.274	4.009	0.025	0.297
diff_year_prior_v2clslavem	-4.915	2.737	0.010	0.184
diff_year_prior_v2clslavef	-4.019	3.731	0.010	0.176
diff_year_prior_v2clstown	-3.472	4.145	0.024	0.240
diff_year_prior_v2clprptym	-3.345	4.414	0.020	0.190
diff_year_prior_v2clprptyw	-3.907	3.834	0.022	0.183

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
diff_year_prior_v2clacjstm	-4.656	2.962	0.015	0.229
diff_year_prior_v2clacjstw	-4.210	2.785	0.016	0.218
diff_year_prior_v2clacjust	-3.102	3.390	0.009	0.185
diff_year_prior_v2clsocgrp	-2.108	2.766	0.013	0.184
diff_year_prior_v2clrgunev	-3.059	2.875	0.002	0.206
diff_year_prior_v2svdomaut	-1.845	4.574	0.013	0.219
diff_year_prior_v2svinlaut	-1.799	4.638	0.016	0.219
diff_year_prior_v2svstterr	-49.200	38.700	0.006	2.347
diff_year_prior_v2cseeorgs	-2.795	4.327	0.026	0.288
diff_year_prior_v2csreprss	-3.985	4.545	0.021	0.315
diff_year_prior_v2cscnsult	-3.010	3.130	0.023	0.300
diff_year_prior_v2csprtcpt	-3.479	4.308	0.029	0.276
diff_year_prior_v2csgender	-2.463	2.908	0.022	0.179
diff_year_prior_v2csantimv	-3.801	3.993	-0.009	0.353
diff_year_prior_v2csrlgrep	-2.723	4.356	0.012	0.256
diff_year_prior_v2csrlgcon	-2.856	3.472	0.014	0.237
diff_year_prior_v2mecenefm	-4.216	4.783	0.021	0.330
diff_year_prior_v2mecrit	-2.574	4.959	0.026	0.301
diff_year_prior_v2merange	-4.102	4.757	0.028	0.293
diff_year_prior_v2meharjrn	-2.438	4.012	0.019	0.286
diff_year_prior_v2meslfcen	-3.382	4.656	0.021	0.315
diff_year_prior_v2mebias	-3.820	4.589	0.027	0.326
diff_year_prior_v2mecorrpt	-2.923	5.011	0.020	0.272
diff_year_prior_v2pepwrses	-2.940	3.437	0.001	0.224
diff_year_prior_v2pepwrsoc	-2.018	3.084	0.012	0.188
diff_year_prior_v2pepwrgen	-2.046	2.760	0.028	0.185
diff_year_prior_v2pepwrort	-1.049	2.210	0.019	0.136
diff_year_prior_v2peedueq	-2.006	2.603	0.007	0.168
diff_year_prior_v2pehealth	-2.106	2.241	0.006	0.168
diff_year_prior_v2x_accountability	-2.082	2.455	0.016	0.149
diff_year_prior_v2x_veracc	-2.361	2.575	0.015	0.250
diff_year_prior_v2x_diagacc	-1.806	2.640	0.017	0.160
diff_year_prior_v2x_horacc	-2.186	2.804	0.012	0.186
diff_year_prior_v2xex_elecleg	-1	1	0.004	0.151
diff_year_prior_v2x_civlib	-0.589	0.676	0.005	0.046
diff_year_prior_v2x_clphy	-0.698	0.735	0.004	0.055
diff_year_prior_v2x_clpol	-0.715	0.828	0.005	0.058
diff_year_prior_v2x_clpriv	-0.464	0.784	0.004	0.042
diff_year_prior_v2x_corr	-0.644	0.486	0.001	0.035
diff_year_prior_v2x_EDcomp_thick	-0.542	0.700	0.005	0.056
diff_year_prior_v2x_elecreg	-1	1	0.005	0.183
diff_year_prior_v2x_freexp	-0.554	0.794	0.005	0.058
diff_year_prior_v2x_gencl	-0.467	0.630	0.004	0.038
diff_year_prior_v2x_gencs	-0.263	0.588	0.006	0.038
diff_year_prior_v2x_hosabort	-1	1	0	0.050

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
diff_year_prior_v2x_hosinter	-1	1	0	0.128
diff_year_prior_v2x_pubcorr	-0.679	0.562	0.001	0.040
diff_year_prior_v2x_rule	-0.486	0.639	0.001	0.040
diff_year_prior_v2xcl_acjst	-0.594	0.688	0.003	0.045
diff_year_prior_v2xcl_disc	-0.690	0.753	0.005	0.060
diff_year_prior_v2xcl_dmove	-0.446	0.827	0.003	0.044
diff_year_prior_v2xcl_prpty	-0.473	0.636	0.005	0.035
diff_year_prior_v2xcl_slave	-0.511	0.708	0.002	0.035
diff_year_prior_v2xcs_ccsi	-0.524	0.767	0.005	0.057
diff_year_prior_v2xel_elecparl	-1	1	-0.003	0.643
diff_year_prior_v2xel_elecpres	-1	1	0.002	0.444
diff_year_prior_v2xex_elecreg	-1	1	0.009	0.157
diff_year_prior_v2xlg_elecreg	-1	1	0.005	0.178
diff_year_prior_v2x_ex_confidence	-1	1	0.002	0.083
diff_year_prior_v2x_ex_direlect	-1	1	0.006	0.158
diff_year_prior_v2x_ex_hereditary	-0.595	0.500	0	0.020
diff_year_prior_v2x_ex_military	-0.732	1	-0.002	0.085
diff_year_prior_v2x_ex_party	-0.800	0.800	-0.001	0.049
diff_year_prior_v2x_execorr	-0.661	0.585	0	0.046
diff_year_prior_v2x_legabort	-1	1	0	0.061
diff_year_prior_v2xlg_leginter	-1	1	0	0.197
diff_year_prior_v2x_neopat	-0.709	0.544	-0.002	0.041
diff_year_prior_v2xnp_client	-0.505	0.297	-0.001	0.040
diff_year_prior_v2xnp_pres	-0.837	0.670	-0.003	0.050
diff_year_prior_v2xnp_regcorr	-0.652	0.464	0.001	0.044
epr_groups	1	58	4.656	5.549
epr_elf	0.013	1	0.578	0.298
epr_excluded_groups_count	0	55	2.416	4.898
epr_excluded_group_pop	0	0.980	0.153	0.215
epr_inpower_groups_count	1	15	2.240	2.138
epr_inpower_groups_pop	0.020	1.044	0.774	0.255
epr_regaut_groups_count	0	42	0.810	3.430
epr_regaut_group_pop	0	1	0.058	0.199
gdp_growth	-83.372	453.824	3.582	10.337
gdp_pc	77.606	117929.820	9906.629	15970.381
gdp_pc_growth	-83.812	444.666	1.771	10.268
gdp_log	18.459	30.460	23.999	2.193
gdp_pc_log	4.352	11.678	8.093	1.552
state_age	1	202	74.118	61.097
pop	4.724	14.159	9.016	1.623

Table 9: Variables and Descriptive Statistics (cont.)

Variable	Min	Max	Mean	SD
pt_coup_attempt	0	1	0.034	0.182
pt_coup_attempt_num	0	4	0.038	0.216
pt_coup_num	0	2	0.018	0.137
pt_coup	0	1	0.018	0.132
pt_failed_coup_attempt_num	0	4	0.020	0.157
pt_failed_coup_attempt	0	1	0.018	0.134
pt_coup_total	0	23	1.977	3.344
pt_coup_attempt_total	0	28	3.088	4.757
pt_coup_num5yrs	0	5	0.100	0.365
pt_coup_attempt_num5yrs	0	9	0.204	0.617
pt_coup_num10yrs	0	6	0.218	0.614
pt_coup_attempt_num10yrs	0	10	0.443	1.051
years_since_last_pt_coup	0	202	42.533	45.415
years_since_last_pt_failed_coup_attempt	0	202	49.187	54.313
years_since_last_pt_coup_attempt	0	202	38.090	45.708
internal_confl	0	1	0.169	0.375
internal_confl_major	0	1	0.051	0.220
internal_confl_minor	0	1	0.130	0.337
internal_confl_part	0	1	0.278	0.448
internal_confl_part_major	0	1	0.141	0.348
internal_confl_part_minor	0	1	0.176	0.381
war	0	1	0.031	0.173
war_major	0	1	0.017	0.131
war_minor	0	1	0.014	0.117
any_conflict	0	1	0.293	0.455
any_conflict_major	0	1	0.154	0.361
any_conflict_minor	0	1	0.183	0.387
ext_conf	0	1	0.156	0.363
ext_conf_major	0	1	0.108	0.310
ext_conf_minor	0	1	0.068	0.252