

# Ensemble forecasting of irregular leadership change

Research and Politics  
October-December 2014: 1–7  
© The Author(s) 2014  
DOI: 10.1177/2053168014557511  
rap.sagepub.com  


Andreas Beger, Cassy L Dorff and Michael D Ward

## Abstract

Using updated Archigos Data, as well as structural and event data, we construct a split-population duration model of irregular leadership changes. These are leadership changes that occur outside of the normal, legal framework for leadership transitions. Our model was estimated in March 2014 and produced probability estimates of leadership changes in many countries in the world. We used a wisdom-of-the-crowds approach to combining estimates from various different models. Ukraine and Thailand are among those in which we had the highest predictions for irregular change of leaders.

## Keywords

Predictions, coups, rebellions, protests

## Introduction

Predicting elections is hard enough. Elections may be the easy case. They come around fairly regularly and generally have set rules for their resolution that are observable.

However, leaders of countries often change for irregular reasons. In late February 2014, pro-Russian President Viktor Yanukovich of Ukraine fled the capital after mass protests erupted into violence, and parliament appointed an interim President to rule until the May elections. A month earlier, in the Central African Republic, Muslim President Michel Djotodia was forced out of office in the face of escalating violence between the Muslim Séléka regime and the largely Christian anti-balaka coalition. In July 2013, the military in Egypt staged a coup and removed democratically elected Mohammed Morsi from the Presidency, following waves of protest against Muslim Brotherhood rule.

Each of these three events – a successful mass protest campaign, a successful rebellion, and a coup d'état – all share the same outcome: the sudden removal of a sitting leader by means outside the “normal” range of political competition. We call this outcome *irregular leadership change* (ILC): the unexpected removal of the principal political leader through means that contravene a state's conventions and laws. Instead of addressing specific mechanisms that drive different types of ILCs, such as narrow conspiracies, mass protests, or armed insurrections, we

instead focus on what leads to the common outcome of ILC. Viewed this way, there have been about four dozen ILCs around the world between 2001 and February 2014, beginning with Mullar Omar's seizure of power in Afghanistan in 2001 and ending with Yanukovich's departure from Ukraine in February of 2014. Only about 40% of these are coups d'état.

## What causes irregular leadership change?

Early interest in coups began in political science during the 1960s (Huntington, 1968; Jackman, 1978; Johnson et al., 1984). Inspired by a wide range of coups that took place in Africa, early work focused on the structural determinants of coups. Goemans and Marinov (2011) note three distinct classes of arguments for why coups occur: political instability resulting from rapid economic modernization (Deutsch, 1961); political illegitimacy following lackluster economic performance and development (McGowan,

Duke University, USA

### Corresponding author:

Michael D Ward, Duke University, 140 Science Drive, 208 Gross Hall, Durham, 27708, USA.

Email: michael.d.ward@duke.edu



Creative Commons CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License (<http://www.creativecommons.org/licenses/by-nc-nd/3.0/>) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<http://www.uk.sagepub.com/aboutus/openaccess.htm>).

2003); and conditions that increase the likelihood of military intervention in politics (Jenkins and Kposowa, 1990; Johnson et al., 1984). These arguments are not disjoint: while one set informs us about the conditions under which a coup might occur, for example as a result of certain structural conditions like the political system, factionalism, or a politicized military, the other provides traction on *when* a coup may occur if the structural conditions are ripe.

The wave of revolutions that brought down communism in Eastern Europe in the early 1990s and the wave of revolutions during the Arab Spring in 2011 have each led scholars to attempt to explain how these revolutions could have so unexpectedly affected monumental change in once “stable” regimes (e.g. Kuran, 1995). Explanations for the suddenness and apparent unpredictability of such revolutions have focused on tipping points that lead to cascades of protest (Kuran, 1991).

Coups are very different from revolutions affected by mass protests, but they have common explanations: conflict between the government and dissidents, bad civil–military relations, and a distinction between general risk and immediate triggers. Empirical work has focused on general risk using static or slow-moving variables such as regime types, economic performance, or military budgets. These might tell us whether there is widespread dissatisfaction with a political system, or troubled civil–military relations and a politicized military, but do not tell us much about the specific timing of events. Triggering events have received less attention, but plausible candidates include indicators sensitive to escalating confrontations between the military or protesters and the reigning government.

ILCs, like coups, revolutions, and rebellions, are apparently heterogeneous processes. The selectorate theory of political survival (Bueno de Mesquita et al., 2005) hints at the possibility of a general framework. It approaches transitions from the leader’s decision-making perspective in which coalition partners must be placated in order to stay in power. Whether it is just a few groups, or a large slice of the population, if those who keep the status quo leaders in power are unhappy, then there is likely to be a turnover. However, even this theory does not draw the distinction between regular and irregular changes.

In short, no one has a unified theoretical explanation of why ILCs occurs.

## Immunity and risk

In modeling ILCs, we address the split between structural risk and immediate triggers. Many approaches to modeling conflicts and abrupt transitions look at all possible cases, and identify variables that can help explain the underlying data. One should not just analyze the cases where there is an ILC (known as selecting on the dependent variable). Yet, we need to gain leverage on the timing of events in those countries that are unstable. We use split-population models

to help with these two issues in studying the irregular and abrupt nature of leadership changes.

Split-population duration models posit that not all cases are at risk of failure. For all practical purposes, countries like Canada are unlikely to experience ILC within our time period of interest, whereas many countries in Africa and the Middle East have experienced ILCs over the past decades. The important point from a modeling perspective is to conceptually separate at-risk countries from those that are practically immune. Once separated, the hazard of an ILC can be better evaluated for all countries.<sup>1</sup>

One advantage of this modeling approach is that it allows covariates to have both a long-term and short-term impact, depending on where they enter the model. Variables that enter the immunity equation have a very long-term impact because they change the probability of being at risk at all. Variables in the duration equation have a short-term impact that modifies the expected duration until the next failure. We use this approach to estimate the duration of regimes over the period from 1955 to the present.

## A strategy for prediction

Rather than engage in testing hypotheses or conducting a horserace of statistical significance (Ward et al., 2010), we instead develop a suite of different models, each capturing different insights. We then see how well these perform on data that we have held back and develop a set of weights that tell us how well each model performs in these calibration data. Then, we use the calibration weights to help construct a probability density that combines all the individual models in the suite into one single estimate, which is then examined with an additional set of data that we have held back for this specific purpose. This general approach has proven useful in a number of areas. It is called Ensemble Bayesian Model Averaging (EBMA; Montgomery et al., 2012; Raftery et al., 2005). Among the more famous demonstrations of this kind of ensemble wisdom was a competition to guess the weight of an ox at the West of England Fat Stock and Poultry Exhibition. Galton (1907) famously demonstrated that, while individual entrants were often wildly inaccurate, aggregating these into an average resulted in a remarkably accurate estimate.

We measure ILCs using the Archigos data on political leaders, which includes the duration of leadership and whether it starts or stops in an irregular fashion. We have updated these data to the present and they provide the duration data we model. These data are combined with our covariates at the monthly level, focusing on the period between 2001 and the present. Our covariates fall into three broad categories. The first are structural variables such as GDP per capita, the Amnesty Political Terror Scale, or regime types. These variables tend to be measured at the country–year level, and mostly vary between countries, but vary less within any particular country. Thus they are more

useful for distinguishing risk sets than predicting the timing of particular events.

The structural variables include several economic and financial indicators such as GDP, population, mortality, military expenditures, broadband subscribers, cell service subscribers, foreign direct investment and consumer price index (CPI) from the World Development Indicators (World Bank Group, 2013), the Polity regime variables (Marshall and Jaggers, 2002), indicators for the number and power relationships of ethnic groups from the Ethnic Power Relations data (Cederman et al., 2009), and the Political Terror Scale (Wood and Gibney, 2010), as well as secondary measures constructed from the Archigos data, such as indicators for leaders who entered irregularly or through foreign imposition (Goemans et al., 2009).

The behavioral variables are constructed from the Integrated Crisis Early Warning System (ICEWS) event data, and record the number of certain types of events in a country over the course of a month, for example protests directed at the government. The ICEWS event data are based on (machine) coded media reports that are parsed for actors, locations, and actions to create event records, using the conflict and mediation event observations (CAMEO) ontology. We include aggregations of events, particularly so-called quad variables that capture verbal and material conflict and cooperation within government and between government and dissidents. For example, verbal cooperation includes making positive public statements, appeals, or consultations, while verbal conflict captures reports of investigations, public demands, or threats. These variables change over time within countries, making them useful for timing the onset of events.

The third set of variables includes spatial lags of the behavioral, event-based variables. A spatial lag captures neighborhood effects, for example the average level of protests in Egypt's neighbors at the time of the uprisings (Ward and Gleditsch, 2008). There are different ways to define what constitutes a country's neighborhood, and we include weights constructed on the basis of the four nearest neighboring countries, the distance between country centroids, and Gower distances (Gower, 1971) of country's similarity on either political, economic, or event measures.

From these variables we construct seven models, each concentrating on substantive themes we believe relevant in understanding political survival.

### Leader characteristics

Drawing on the literature on leadership tenure (Acemoglu and Robinson, 2006; Bueno de Mesquita et al., 2005; Svobik, 2012), we build a model that captures the leaders' individual characteristics, as well as internal regime cooperation. The literature on leadership survival focuses not only on a leader's ability to consolidate power over time, but also considers that as a leader consolidates power, they

are more likely to create discontent among those who are not politically represented by the regime. The risk equation thus includes a count of the months a leader has been in power. To capture the legitimacy of a leader and by association his or her government, we include two further variables in the risk equation that indicate whether the current leader of a state entered power through irregular means or by foreign imposition. Leaders who entered through illegitimate, irregular means might themselves be more likely to suffer the same fate. The duration equation uses the material behavior of dissidents, whether cooperative or conflictual, to capture the timing. We use material rather than verbal actions to model the timing of an ILC against illegitimate leaders.

### Public discontent

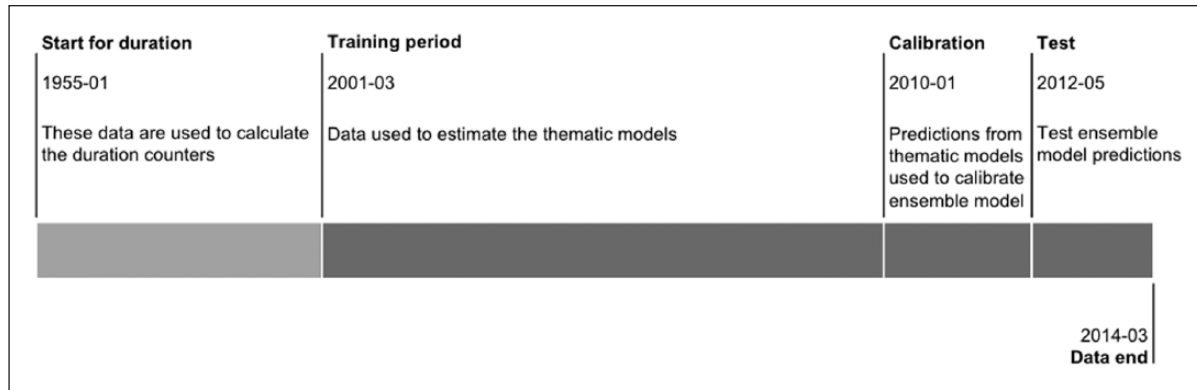
The public discontent model focuses on verbal interactions as well as protests to provide an early warning indicator of ILCs. We also examine verbal cooperation within government, primarily but not exclusively as an indicator of the health of civil–military relations. Since the level of public, verbal interactions in a society is related to access to media and the ability to voice demands, for the model in the risk equation we include per capita measures of Internet users and cell subscribers. Many authoritarian governments implement censorship to control the information available to citizens. We also include the fraction of excluded population in a country as a control, since minority governments facing a large opposition have strong incentives to display unity.

### Global instability

Our third model is loosely based on the main components of the Goldstone et al. (2010) model. Using these findings on what factors drive global instability, we have created a model loosely based on theirs, but necessarily different given our different modeling strategy and data resolution. In our version, the partial democracy with factionalism indicator did not perform as well as simply including the Polity participation of competitiveness variable, which captures whether “alternative preferences for policy and leadership can be pursued in the political arena.” Echoing the Goldstone approach, we include GDP and the percentage of the population excluded from the political process into the risk equation. Then, to predict the timing of ILC, we include participation competitiveness, a measure of conflict within the four nearest neighbors, as well an indicator of female life expectancy at birth.

### Anti-regime protests

This thematic model is entirely focused on protest. Civil resistance campaigns are an effective means for achieving



**Figure 1.** Data partitions.

leadership change. The literature on both coup-proofing (Pilster and Böhmelt, 2011; Quinlivan, 1999) and civil resistance campaigns (Chenoweth and Stephan, 2011) describe a key force behind protest movements: their ability to influence the military. A pivotal movement in many civil resistance campaigns is the moment when state forces stop obeying orders from the head of state, and refuse to openly repress protestors. This model captures the basic intuition of this argument by including slower moving structural variables, such as low levels of domestic crises and military expenditure, into the risk equation. This model is structured by the argument that the least satisfied militaries will be most likely to resist commands to repress. In the duration equation we account for protest and conflict in different forms: ethnic-religious violence, rebellion, protest events, and nearby rebellion events in other countries.

## Contagion

This model captures the concept of conflict contagion. To model the risk for successful contagion of mass protests or other conflict that may lead to an ILC, we include the country's Amnesty International Political Terror Scale value, which captures overall repressiveness, as well as opposition resistance, which counts the number of events conducted by groups associated with armed anti-government groups. The latter largely varies between rather than within countries, and we thus include it as a static variable. These two variables capture the overall security climate in a country. To further refine the general risk posed in a repressive society with ongoing terror or political violence, we include an indicator of temporally proximate elections. This variable identifies whether an election will occur in the near future or has occurred in the near past. Finally, we include the country's population size as an indicator of society's inertia and resistance to outside influences.

Having specified risk, we use two spatial weights of opposition resistance and state repression in neighboring countries to model the timing until contagion, and hence increased chance of ILC, occurs.

## Internal conflict

The internal conflict model uses GDP per capita, the proximity of the next national election, and the level of Autocracy in the country as general indicators of risk, while focusing on intra-governmental conflict and the widespread use of cell technology as duration triggers. Intra-governmental tensions, protests to the government, and the number of cell phones are taken to interact to influence the duration of leadership tenure and the likelihood of an irregular transfer. First-order components of this interaction also are included in the duration equation, but the second-order interactions (e.g. the two-way interactions) are excluded as they cause instabilities in the likelihood.

## Financial risks

This model assumes that financial instability may unseat leaders who are already in a precarious situation. The baseline risk is determined by GDP per capita, as a measure of general prosperity and the looming presence of the next election, as well as the size of the country as measured by population. In addition, it includes the Amnesty assessment of terrorism (stability) and the degree of anti-government. If a country is in the high risk set, it is the degree of inflation, as measured by consumer prices, and the health of the country's international financial reserves (taken from the International Monetary Fund's (IMF's) International Financial Statistics (IFS)) that affect most directly the duration of leadership.

## Summary of modeling strategy

Prior to statistical estimation, we divide our data into separate partitions. A fourth partition is the data yet to be observed: that is, the future. We use this tripartite division to guard against overfitting (see Figure 1).

Each theme is estimated separately in the training data using a split-population estimator that we have created. The seven streams of predictions from these models for the



**Table 1.** Comparing the ensemble model to two simple baseline approaches.

Model	In-sample		Out-of-sample	
	AUC	F	AUC	F
Ensemble	0.856	0.068	0.839	0.114
Logit	0.951	0.286	0.776	0.059
Split-dur.	0.744	0.085	0.664	0.080

F: harmonic mean of precision/recall; AUC: area under the receiver operating characteristic curve.

calibration period are examined with the EBMA approach to calibrate a set of performance weights. These weights and the underlying theme prediction combine to form the ensemble. Finally, we examine the performance of each theme as well as the ensemble in a final partition of the data, the test partition. In this short presentation, we spare the reader all of these details, and turn to the actual forecasts made by this approach.<sup>2</sup>

At the request of several readers of this work and in honor of Christopher Achen (2005), we ran a “garbage can” logit regression with the 39 covariates from our theme models. The findings are not surprising. Such models typically have the characteristic of over-fitting the in-sample data. Much social science seems to stop there and declare victory. However, such garbage can models almost always are terrible at out-of-sample predictions. That is exactly what we found: our theme models (as well as the ensemble composite) are about one-third better in terms of recall and precision. We could not estimate a similar all-in split-duration model due to convergence issues, but a version that includes the 13 covariates from the two models that receive the largest weights in the ensemble, contagion and internal conflict, also performs worse out of sample. Table 1 illustrates this fit of the two baseline models versus our ensemble model.

## Examining predicted change

Using the ensemble model and data from March 2014 we create forecasts for the probability of ILC over the period from April to September 2014. We aggregate the monthly forecasts produced by this model to an overall probability of ILC anytime during this time period, and Table 2 shows the 10 highest forecasts that result.<sup>3</sup>

These predictions were made in March 2014. Probabilities are not certainties. For every 20 estimates that there is a 0.05 chance of rain; for a properly calibrated model, one should expect it to rain at least once. ILCs are very rare events. Our top five predictions include Ukraine, Bosnia and Herzegovina, Yemen, Egypt, and Thailand.

Ukraine lost the Crimea to Russia this spring, but the protests in the winter of 2013 continued into the new year and violent protests occurred in the middle of February 2014 that were in part a response to the so-called Anti-Protest

**Table 2.** Top 10 forecasts for irregular leadership change between April and September 2014 (six months) using March 2014 data.

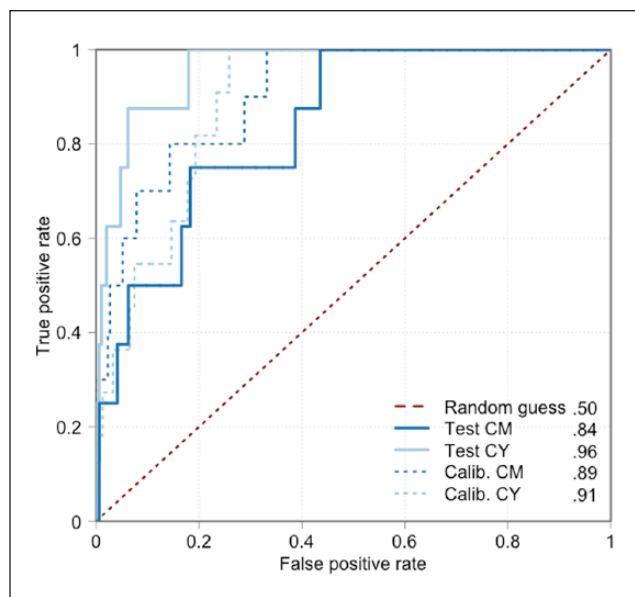
Country	Probability
Ukraine	0.28
Bosnia and Herzegovina	0.19
Yemen	0.10
Egypt	0.07
Thailand	0.06
Guinea	0.05
India	0.04
Turkey	0.04
Libya	0.03
Central African Republic	0.03

laws enacted in the previous month. By the end of February, Parliament essentially ousted the President and scheduled a May 25 election. This created a succession crisis in which the deposed president, Yanukovich, and his supporters in Russia began to create a larger conflict in Ukraine. Russian involvement spawned further local conflict within Ukraine. Continuing conflict within the government and separatist activities in the east, in combination with a new government, place Ukraine at the top of our model predictions. Poroshenko won the May 2014 elections, but several thousand have been killed and many more have fled the violence in Eastern and Southern Ukraine.

Bosnia and Herzegovina has been the site of substantial and widespread anti-government protests in early 2014, the so-called Bosnian Spring. These were organized in large part because of the frailty of the economy, the high level of unemployment, and the non-payment of pensions. Prime Minister Vjekoslav Bevanda has minimized the protests. However, as the leader of an increasingly weak central government, the greatest instability may reside in the locally governed regions.

Yemen is the site of protests accompanied by the presence of a very powerful Al-Qaeda army. That, alongside a set of rulers widely reported to be corrupt, creates an unstable situation. Yemen is in the throes of (another) reorganization in which central authority seems likely to devolve to regional ruling coalitions.

Egypt sees an outbreak of protests and violence every year in February to celebrate the resignation of Mubarak and the start of the so-called revolution in the early spring of 2011. In mid-2013 these protests spread and sparked a coup d'état that displaced Morsi. In mid-2013 Mansour was appointed as acting president. In early 2014 a new constitution was overwhelmingly ratified by Egyptian voters, even though roughly two-thirds of the potential voters avoided participating. In May 2014, a presidential election was won by Abdel Fattah el-Sisi, the Egyptian commander-in-chief. Tension remains high and the legitimacy and popularity of the current regime is tenuous, at best.



**Figure 2.** Forecast model receiver operating characteristic curves. (Color online only.)

Thailand has been a puzzling cauldron of political conflict for over a decade. Thaksin Shinawatra was overthrown by a coup d'état in the fall of 2006, and the resulting junta instituted martial law and forbade many political activities until mid-2007. Things continued to be contentious and violent, but by mid-2011 things had calmed down and Yingluck Shinawatra handily won the election. Toward the end of 2013, protests heated up quite a bit as did demands for the resignation of Yingluck. Scheduled elections were not held in 2014, as the listing of candidates outraged anti-government forces. As this was written, on May 7, the (Supreme) Constitutional court ruled that Yingluck had abused power and was to be removed from the Prime Ministership. It is unclear how this will turn out, except to note that this will constitute another ILC in mid-2014.

## Model fit statistics

Any probabilistic model for ILCs will have a tradeoff between false positives and false negatives. In a random guess this tradeoff is even, leading to the receiver operating characteristic (ROC) curve shown in red in Figure 2. Any useful model should exceed it. The ROC curves for the ensemble – in-sample (dashed) and out-of-sample (solid) – show good fit with areas under the curve (AUCs) above 0.84 for our monthly data and above 0.91 when we annualize our predictions for comparison with work at the country-year.

With infrequent events such as ILCs it also makes sense to evaluate recall and precision. Recall is the fraction of events accurately predicted by a model, and precision is how many positive predictions turn out to be actual events, or how believable the predictions are. As before, there is a

tradeoff between the two depending on the cutoff one chooses for separating probabilistic predictions to 0/1 values. The ensemble obtains a recall of 0.5 with a precision of 0.016 or 1 in 60 in the monthly data, and around 1 in 5 in annualized data. With data as sparse as ours, a model needs to perform extremely well by conventional standards in order to make predictions we can take at face value.

## Conclusion

We used new, temporally disaggregated data that included behavioral variables derived from event data. We also employed split-population duration and ensemble modeling approaches to examine ILCs over the period from 1955 to the present. Each of these aspects is novel in the study of leadership change. In so doing we also developed a suite of new empirical models that were measured monthly. In addition, we then combined the forecasts of each of these empirical models using EBMA to produce a single probability estimate that benefits from the so-called “wisdom of the crowds.” Along the way, we updated the dependent variable over the past two and a half years.

The suite of models we developed is examined in historical training data and was evaluated in test data that were not used in the initial data construction. In both contexts, the models are accurate and well calibrated. Finally, we use a weighted ensemble combination of these models to produce six-month forecasts of the conditional hazard over the period from April 2014 until September 2014. These predictions are discussed above, but in summary seem plausible. Indeed, two of the top 10 forecasts are Ukraine and Thailand, both currently in the throes of transition crises.

In our attempt to forecast ILCs, we have created a complex framework that breaks with many conventions in previous scholarship. This opens us to many criticisms, such as questions about the rigor of the thematic models in our suite, or the utility of the ILC concept. We tried to root the models in themes notable in the literature on regime change. However, our approach is inherently modular and open to the inclusion of other models with the ensemble as arbiter of their usefulness in prediction. This study may serve as a foundation for future inquiry and to encourage scholars to conduct similar investigations at the country-month level.

Many months pass in each country without an ILC. They are rare. Our modeling approach has the goal to accurately forecast ILCs, and the rarity of these events has led to the novel aspects we have presented here. Still, we are looking for needles in a haystack. Even our 10 highest predictions have low probabilities of ILC. However, someone once said, “reality is a low probability event.”

## Funding

This work was supported by the Political Instability Task Force (PITF). The PITF is funded by the Central Intelligence Agency. The views expressed in this report are the authors alone and do not represent the views of the US Government.

## Declaration of conflicting interest

The authors declare that there is no conflict of interest.

## Notes

- 1 The likelihood is given as a product of the immunity and the risk:  

$$\mathcal{L}\{\theta | (t_1, \dots, t_n)\} = \prod_{i=1}^N \{(1 - \pi)f(t_i)\}^{\delta_i} \times \{\pi + (1 - \pi)S(t_i)\}^{1 - \delta_i}.$$
- 2 Analysis was done in R 3.0.2 using R packages that are either available on CRAN or included in the replication files available at our dataverse: <http://dx.doi.org/10.7910/DVN/27482>. More, many more, details are available in the full report, available at <http://arxiv.org/abs/1409.7105>.
- 3 Using  $p^* = 1 - \prod_{t=1}^6 (1 - p_t)$ , where  $p_t$  is the forecast  $t$  months ahead.

## References

- Acemoglu D and Robinson JA (2006) *Persistence of power, elites and institutions*. Technical report, National Bureau of Economic Research.
- Achen CH (2005) Let's put garbage-can regressions and garbage-can probits where they belong. *Conflict Management and Peace Science* 22(4): 327–339.
- Bueno de Mesquita B, Smith A, Siverson RM, et al. (2005) *The Logic of Political Survival*. Cambridge, MA: MIT Press.
- Cederman L-E, Min B and Wimmer A (2009) Ethnic power relations dataset. Available at: <http://www.epr.ucla.edu/> (accessed 1 May 2009).
- Chenoweth E and Stephan MJ (2011) *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. New York: Columbia University Press.
- Deutsch KW (1961) Social mobilization and political development. *The American Political Science Review* 55(3): 493–514.
- Galton F (1907) Vox populi. *Nature* 75(1949): 450–451.
- Goemans HE, Gleditsch KS and Chiozza G. (2009) Introducing Archigos: A dataset of political leaders. *Journal of Peace Research* 46(2): 269–283.
- Goemans HE and Marinov N (2011) Elections after the coup d'état: The international community and the seizure of executive power. Unpublished manuscript.
- Goldstone JA, Bates RH, Epstein DL, et al. (2010) A global model for forecasting political instability. *American Journal of Political Science* 54(1): 190–208.
- Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* 27(4): 857–871.
- Huntington SP (1968) *Political Order in Changing Societies*. New Haven, Ct: Yale University Press.
- Jackman RW (1978) The predictability of coups d'état: A model with African data. *The American Political Science Review* 72(4): 1262–1275.
- Jenkins JC and Kposowa AJ (1990) Explaining military coups d'état: black Africa, 1957–1984. *American Sociological Review* 55(6): 861–875.
- Johnson TH, Slater RO and McGowan P (1984) Explaining African military coups d'état, 1960–1982. *The American Political Science Review* 78(3): 622–640.
- Kuran T (1991) Now out of never: The element of surprise in the East European revolution of 1989. *World Politics* 44(01): 7–48.
- Kuran T (1995) The inevitability of future revolutionary surprises. *American Journal of Sociology* 100(6): 1528–1551.
- Marshall MG and Jaggers K (2002) Polity IV project: Political regime characteristics and transitions, 1800–2002. Available at: <http://www.systemicpeace.org/polity/polity4.htm>
- McGowan PJ (2003) African military coups d'état, 1956–2001: Frequency, trends and distribution. *The Journal of Modern African Studies* 41(3): 339–370.
- Montgomery JM, Hollenbach FM and Ward MD (2012) Improving predictions using ensemble Bayesian model averaging. *Political Analysis* 20(3): 271–291.
- Pilster U and Böhmelt T (2011) Coup-proofing and military effectiveness in interstate wars, 1967–99. *Conflict Management and Peace Science* 28(4): 331–350.
- Quinlivan JT (1999) Coup-proofing: Its practice and consequences in the Middle East. *International Security* 24(2): 131–165.
- Raftery AE, Gneiting T, Balabdaoui F, et al. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review* 133(5): 1155–1174.
- Svolik MW (2012) *The Politics of Authoritarian Rule*. New York: Cambridge University Press.
- Ward MD and Gleditsch KS (2008) *Spatial Regression Models*. Vol. 155. Beverly Hills, Ca: Sage.
- Ward MD, Greenhill BD and Bakke KM (2010) The perils of policy by  $p$ -value: Predicting civil conflicts. *Journal of Peace Research* 47(4): 363–375.
- Wood RM and Gibney M (2010) The Political Terror Scale (PTS): A re-introduction and a comparison to CIRI. *Human Rights Quarterly* 32(2): 367–400.
- World Bank Group (2013) *World Development Indicators 2013*. Washington DC: World Bank Publications.