Appendix for
"Lessons from near real-time forecasting of irregular leadership changes," *Journal of Peace Research*

Michael D. Ward and Andreas Beger

2 November 2016

Contents

1 Comparison of EBMA ensemble, LASSO, and random forrest

We conducted a comparison of the ensemble with a random forest and lasso regression as alternatives. We use the superset of the split-population training data and EBMA calibration data as training data, and retained the same test set for out-of-sample testing.

We did not include a kitchen sink model in our comparison. The seven theme models in the ensemble are based on 37 unique variables, while the training data have 84 events. Estimating a model with a ratio of 2.3 events per variable is problematic Vittinghoff and McCulloch (2007). Some of the 37 unique covariates also induce multicollinearity. A full kitchen sink model will estimate, but with missing (NA) estimates for some variables, and prediction will produce rank-deficient fit warnings in R. In light of these problems we chose lasso regression, which performs variable selection as well as regularization, as a more sensible alternative. Both the lasso

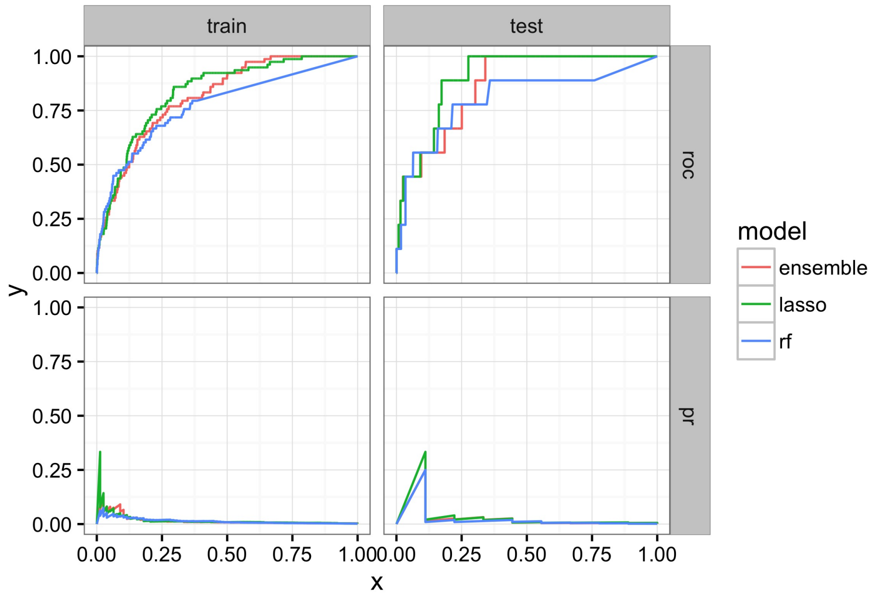

Figure A1

regression and random forest are based on the same set of 37 unique covariates that go into our ensemble forecast.

For the random forest, we used 5,000 trees and the recommended default values for other parameters, although alternative choices do not seem to result in drastic changes. We use the out of bag (OOB) predictions for the in-sample training data, i.e. predictions for the training set observations are calculated on the basis of all trees that did not include that particular case in their randomly-sampled subset of the training data.

The in and out-of-sample fit of the models is summarized in Table A1. Figure A1 also shows the corresponding ROC and precision-recall curves, and separation plots are in Figure A2. Note that the EBMA fit here is different from the fit we report in the article itself–this is because (1) here we combine the main article training and calibration data to use a combined training set and (2) for the tests, we look at monthly predictions, rather than the 6-month test forecasts we conduct in the article.

Table A1: Relative model fit

Data	Model	AUC-PR	AUC-ROC	Brier
Train	Ensemble	0.016	0.816	0.002
Train	Lasso	0.017	0.836	0.002
Train	Random Forest	0.014	0.772	0.002
Test	Ensemble	0.024	0.864	0.001
Test	Lasso	0.031	0.900	0.002
Test	Random Forest	0.022	0.805	0.002

The lasso regression outperforms the ensemble by a small but noticeable margin, and both in turn perform better than the random forrest. The separation plots show the same pattern. The results match Jay Ulfelder's finding in his own coup forecast that random forest does not con-

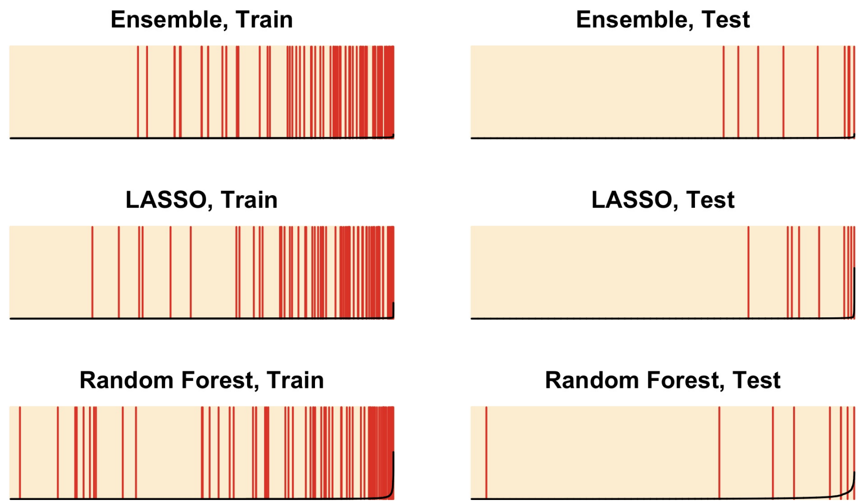

Figure A2: Separation plots for the ensemble, random forest, and LASSO predictive models.

tribute dramatic accuracy improvements over conventional logistic regression,[1] and also those in Muchlinski et al. (2016), who compare random forest to logistic regression for the prediction of rare events.

What are the implications of these results for our project? Lasso and the ensemble are comparable in terms of performance, but the lasso is arguably a less complex method overall. Since the predictions from lasso regression at the core come down to a single linear equation, interpretation and results could be presented similar to what we do with our ensemble now. But since performance is similar, it does not seem worth the cost of switching over. Another drawback is the loss of control over variable selection and thus model specification.

The random forest does not predict as well and also has the drawback of more complex interpretation. While tools like variable importance and partial dependence plots have been developed to unravel the random forest "black box", the tools we are aware of are limited in their ability to aid in the interpretation of model predictions for particular cases, rather than average, model-level general trends.

The partial dependence plot for example is based on evaluating the random forest predictions over a range of values for a variable while holding all other variables at their observed values and averaging over the resulting spread of predictions. The actual association between a variable and the model predictions for a particular case may actually be completely different from the average impact visualized in the partial dependence plot however Goldstein et al. (2015).

In our understanding, there is no way to provide simple heuristics, like "P is high because x is high" with implied corollaries like "P would be lower if x was lower" with a random forest without simulation over values of x and access to the full set of 5,000 trees and relevant data for a case, nor would heuristics thus developed hold for other cases. This is a serious limitation for our purposes.

[1] https://dartthrowingchimp.wordpress.com/2015/01/17/statistical-assessments-of-coup-risk-for-2015/

2 Irregular leadership changes (ILC)

2.1 Definition of "irregular" transition

Irregular leadership changes occur when a leader leaves office or his/her successor enters office in a manner that does not follow a regime's legal or otherwise established conventions for leadership transitions, or when a leader leaves office in a way that procedurally followed that regime's conventions, but did so under coercion from outside forces, e.g. mass protests or the military. Leader refers to the effective leader of a country, the single person who has the largest amount of power over a country's political affairs. A regime's legal or otherwise established conventions can be based on written constitutions or documents, precedent, or otherwise reasonable expectations about how transitions will occur.

Our definition of irregular leadership change (ILC) is a direct adaption of the concepts of "irregular" leadership entry (gain of office) and exit (loss of office) in the Archigos dataset (Goemans, Gleditsch and Chiozza 2009). From the Archigos codebook (1–2):[2]

> Archigos codes the manner in which transfers between rulers occur. Our main interest is whether transfers of power between leaders take place in a regular or irregular fashion. We code transfers as regular or irregular depending on the political institutions and selection mechanisms in place. We identify whether leaders are selected into and leave political office in a manner prescribed by either explicit rules or established conventions. In a democracy, a leader may come to power through direct election or establishing a sufficient coalition of representatives in the legislature. Although leaders may not be elected or selected in particularly competitive processes, many autocracies have similar implicit or explicit rules for transfers of executive power. Leader changes that occur through designation by an outgoing leader, hereditary succession in a monarchy, and appointment by the central committee of a ruling party would all be considered regular transfers of power from one leader to another in an autocratic regime."

We operationalize ILCs using the Archigos data and code an ILC as having occurred when either an irregular exit, irregular entry, or both occur in a given country in a given month.

2.2 Archigos entry and exit coding scheme

The Archigos codebook (2–3) gives more details on the entry and exit coding. Entries are coded as follows, with no further details outside of the relevant textual summaries for each country in the codebook documentation.

0	Leader reached power through regular means
1	Leader reached power through irregular means
2	Leader directly imposed by another state

For leader exits, a two-level typology is used. First, exit is broadly classified into similar categories to entry:

[2]Archigos Codebook: http://www.rochester.edu/college/faculty/hgoemans/Archigos.2.9-August.pdf

-888	Leader still in power
1	Leader lost power through regular means
2	Leader died of natural causes while in power
2.1	Leader retired due to ill health
2.2	Leader lost office as a result of suicide
3	Leader lost power through irregular means
4	Leader deposed by another state

Then, for irregular exists and leaders deposed by another state (exit 3 and 4), an additional exit detail code is given.

0	Leader lost power in a regular manner
1	Leader lost power as a result of domestic popular protest with foreign support
2	. . . without foreign support
3	Leader removed by domestic rebel forces with foreign support
4	. . . without foreign support
5	Leader removed by domestic military actors with foreign support
6	. . . without foreign support
7	Leader removed by other domestic government actors with foreign support
8	. . . without foreign support
9	Leader removed through the threat or use of foreign force
11	Leader removed through assassination by unsupported individual
16	Leader removed in a power struggle within military, short of coup, i.e. without chang- ing institutional features such as a military council or junta
111	Leader removed in an irregular manner through other means or processes

The exit detail codes 9 and 11 are only used for leaders overthrown by other states (exit 4), the rest pertain only to irregular exits (exit 3). The empirical relationship between these two coding levels, using a version of the Archigos data that we updated for more recent time periods, is as follows.

	Exit detail code										
Exit code	2	3	4	5	6	7	8	9	11	16	111
3	32	8	23	5	194	3	23	0	11	21	9
4	0	0	0	0	0	0	0	17	0	0	2

2.3 Our own Archigos updates

The original Archigos data code leaders through 2004. We received a version updated to 2011 from one of the project authors, and have reviewed as well as continued updating the data ourselves, using the Archigos codebook. Our coding has been reviewed by a non-academic audience, and we made several changes in response to those comments. We have also shared it with the original Archigos project investigators.

The text notes that ILCs are coded based on whether either an irregular exit or entry to office took place. Although these two often co-occur, e.g. a coup against a sitting leader is coded as irregular exit for the deposed leader followed by irregular entry for the new leader, they do not always, as for example when a leader is forcibly removed from office but his or her replacement follows succession rules, e.g. a vice president, or when a coup occurs after the death of a sitting leader.

2.4 List of ILCs from 1991 on

Table A2 lists all ILCs from 1991 on. It includes indicators for whether and ILC occurred due to the irregular exit or entry of a leader, or both. Due to coding convention, we list the leader who was in power when the ILC occurred, i.e. the first leader in a transition, even if the transition of irregular because an irregular entry without prior irregular exit occurred.

Only two cases involve an irregular entry without irregular exit for the predecessor: 2007 in Fiji where Senilagakali resigned as prime minster and was followed by Bainimarama, who a year earlier had staged a military coup, and 2011 in Guinea where a military coup occurred after Conte died in office.

3 Model discussion and estimates

3.1 Ensemble and thematic models

Two main considerations drive our modeling framework. First, ILCs are multi-causal phenomena with a multitude of proposed or plausible models for various relevant aspects. As we are interested in prediction, a single model or even an attempt at a single general model does not predict as well as an ensemble of multiple approaches.

The second consideration is that ILCs, like many conflict outcomes, are rare events. There have been fewer than a hundred ILCs over the past 25 years, and the data are even more sparse at the monthly level than they would be with more usual annual data. The base rate of ILCs in our data is 17 events per 10,000 country-months. In comparison, the data in Fearon and Laitin (2003) have a base rate of 167 civil war *onsets* per 10,000 country-years. The rate of civil war *occurrence* is naturally higher still. Although rare events and sparse data are not unusual at all in conflict research, the underlying problem is more difficult than it might be with higher frequency outcomes.

One approach to both of these problems would be algorithmic machine learning solutions which have been used for more challenging but similar problems like document retrieval for search queries. Examples include neural nets, SVM, and decision trees/random forests, which can incorporate nonlinear relationships and have been created specifically for the task of prediction. We shy away from these approaches as they typically are "black box", meaning that the mapping from inputs to outputs is complex and usually cannot be easily interpreted. This conflicts with the premium in forecasting for being able to explain which factors are driving high forecasts.

We instead rely on an ensemble of several distinct models to produce our forecasts, and specifically use Ensemble Bayesian Model Averaging to transform and weight component forecasts from several thematic models (Raftery et al. 2005, Montgomery, Hollenbach and Ward 2012). This approach allows us to aggregate multiple distinct models, each of which captures and predicts some aspect of ILCs, in a principled way to maximize predictive performance.

The EBMA ensemble forecast for forecasts f_k' from k input models is a weighted average of transformed raw input probabilities using estimated weights W that maximize fit over a sample of calibration data. The raw forecast probabilities first undergo an adjustment for bias reduction, $h()$, which transforms the probabilities to the logit scale and pulls them closer to 0 in order to reduce the effect of extreme predictions:[3]

$$
\begin{aligned}
f_k &= h(f'_k) & (1)\\
&= [(1+\text{logit}(|f'_k|))^{1/b}-1]\times[-I(f'_k<1/2)] & (2)
\end{aligned}
$$

where b denotes the extent of compression, and where $I()$ is an indicator function. The logit-scale forecasts then undergo an affine transformation based on parameters from a logistic regression on observed outcomes in the calibration sample before being combined into the final weighted average:

$$
\begin{aligned}
g_k(y|f_k) &= \text{logit}^{-1}(a_{k0}+a_{k1}\times f_k) & (3)\\
p &= \sum_{k=1}^{K} w_k g_k(y|f_k) & (4)
\end{aligned}
$$

To address the rare events problem we draw on the pattern that many countries in our data will for and use split-population duration regressions as the workhorse of the underlying thematic models (for an application in political science, see Svolik 2008).

Split-population duration regression is a mixture of two equations, one that reflects a traditional duration model with time-varying covariates and accelerated failure time form, and a second logistic equation that simultaneously estimates the risk that a case will be at risk of failure at all. The likelihood is given as a product of the immunity π and the risk, where δ_i indicates whether a spell ended in failure:

$$
\begin{aligned}
\mathscr{L}\{\theta|(t_1,\ldots,t_n)\} &= \prod_{i=1}^{N}\left\{(1-\pi)f(t_i)\right\}^{\delta_i}\times\\
&\quad \{\pi+(1-\pi)S(t_i)\}^{1-\delta_i}
\end{aligned}
$$

Practically, the data are grouped into spells that consist of observations for a particular country over time until either ILC occurs, when a new spell starts, or until the right-censoring time. Risk is back-coded for all spells that had an observed ILC, and thus the risk variable is much more extensive than the number of observed ILC country-months itself. This allows the logistic equation to separate country-months that are likely to be in the risk or immune set. The duration equation can accommodate a Weibull or log-logistic form for the hazard function, but the Weibull shape appears to fit better in our thematic models.

[3] The notation is largely consistent with (Montgomery, Hollenbach and Ward 2012) save for the use of terms to denote the transformation equations.

Table A2: List of Irregular leadership changes from 1991 to 2014

	Country	Date	Leader	Irr. Exit	Irr. Entry	Yrs. in power
1	Somalia	1991-01	Siad Barre	1	0	21
2	Thailand	1991-02	Choonhavan	1	0	2
3	Mali	1991-03	Traore	1	1	22
4	Ethiopia	1991-05	Mengistu Marriam	1	1	14
5	Lesotho	1991-05	Lekhanya	1	1	5
6	Russia (Soviet Union)	1991-08	Gorbachev	1	0	6
7	Haiti	1991-09	Aristide	1	1	1
8	Algeria	1992-01	Benjedid	1	1	13
9	Georgia	1992-01	Gamsakhurdia	1	1	1
10	Afghanistan	1992-04	Najibullah	1	1	6
11	Sierra Leone	1992-04	Momoh	1	1	6
12	Azerbaijan	1992-05	Mamedov	0	1	<1
13	Algeria	1992-06	Boudiaf	1	0	<1
14	Azerbaijan	1992-06	Gambarov	0	1	<1
15	Tajikistan	1992-09	Nabiyev	1	0	1
16	Tajikistan	1992-11	Iskandrov	0	1	<1
17	Pakistan	1993-04	Sharif	1	1	2
18	Guatemala	1993-05	Serrano Elias	1	0	2
19	Sri Lanka (Ceylon)	1993-05	Premadasa	1	0	4
20	Azerbaijan	1993-06	Abulfaz Elchibey	1	0	1
21	Guatemala	1993-06	Espina Salguero	0	1	<1
22	Nigeria	1993-08	Babangida	0	1	8
23	Burundi	1993-10	Ndadaye	1	1	<1
24	Nigeria	1993-11	Shonekan	1	1	<1
25	Burundi	1994-04	Ntarymira	1	0	<1
26	Rwanda	1994-04	Habyarimana	1	1	21
27	Gambia	1994-07	Jawara	1	1	29
28	Lesotho	1994-08	Mokhehle	1	1	1
29	Lesotho	1994-09	Letsie III	0	1	<1
30	Solomon Islands	1994-10	Hilly	1	1	1
31	Qatar	1995-06	Khalifah Ath-Thani	1	1	23
32	Comoros	1995-09	Djohar	1	1	6
33	Israel	1995-11	Rabin	1	0	3
34	Niger	1996-01	Ousmane	1	1	3
35	Sierra Leone	1996-01	Strasser	1	1	4
36	Burundi	1996-07	Ntibantunganya	1	1	2
37	Afghanistan	1996-09	Burhanuddin Rabbani	1	1	4
38	Pakistan	1996-11	Bhutto Benazir	1	1	3
39	Congo, Dem. Rep.	1997-05	Mobutu	1	1	32
40	Sierra Leone	1997-05	Kabbah	1	1	1
41	Turkey (Ottoman Empire)	1997-06	Erbakan	1	1	1
42	Cambodia (Kampuchea)	1997-07	Ranariddh	1	1	4
43	Congo	1997-10	Lissouba	1	1	5
44	Comoros	1999-04	Massounde	1	1	<1
45	Niger	1999-04	Mainassara	1	1	3
46	Guinea-Bissau	1999-05	Vieira	1	1	18
47	Pakistan	1999-10	Sharif	1	1	3
48	Cote D'Ivoire	1999-12	Konan Bedie	1	1	6

continued on next page

continued from previous page, 2000–2014

	Country	Date	Leader	Irr. Exit	Irr. Entry	Yrs. in power
49	Ecuador	2000-01	Mahuad	1	1	1
50	Fiji	2000-05	Chaudhry	1	1	1
51	Solomon Islands	2000-06	Ulufa'alu	1	0	3
52	Fiji	2000-07	Bainimarama	1	0	<1
53	Cote D'Ivoire	2000-10	Guei	1	1	1
54	Congo, DRC	2001-01	Laurent Kabila	1	0	4
55	Afghanistan	2001-11	Mullah Omar	1	0	5
56	Madagascar	2002-07	Ratsiraka	1	1	5
57	Central African Republic	2003-03	Patasse	1	0	9
58	Guinea-Bissau	2003-09	Kumba Iala	1	1	4
59	Georgia	2003-11	Shevardnadze	1	0	12
60	Haiti	2004-02	Aristide	1	0	3
61	Kyrgyz Republic	2005-04	Akayev	1	0	14
62	Bolivia	2005-06	Carlos Mesa	1	0	2
63	Mauritania	2005-08	Sidi Ahmed Taya	1	1	21
64	Nepal	2006-04	Gyanendra	1	0	1
65	Solomon Islands	2006-05	Rini	1	0	<1
66	Thailand	2006-09	Thaksin Shinawatra	1	0	6
67	Fiji	2006-12	Laisenia Qarase	1	0	6
68	Bangladesh	2007-01	Iajuddin	1	0	<1
69	Fiji	2007-01	Senilagakali	0	1	<1
70	Georgia	2007-11	Saakashvili	1	0	4
71	Lebanon	2008-05	Siniora	1	0	<1
72	Mauritania	2008-08	Ould Cheikh Abdellahi	1	1	1
73	Guinea	2008-12	Conte	0	1	25
74	Guinea-Bissau	2009-03	Vieira	1	0	3
75	Madagascar	2009-03	Marc Ravalomanana	1	1	7
76	Honduras	2009-06	Zelaya	1	1	3
77	Guinea	2009-12	Dadis Camara	1	1	1
78	Niger	2010-02	Mamadou	1	1	10
79	Kyrgyz Republic	2010-04	Bakiyev	1	0	5
80	Tunisia	2011-01	Zine Al-Abidine Ben Ali	1	0	23
81	Egypt	2011-02	Mubarak	1	0	29
82	Cote D?Ívoire	2011-04	Laurent Gbagbo	1	1	10
83	Libya	2011-08	Qaddafi	1	1	42
84	Mali	2012-03	Amadou Toure	1	1	10
85	Guinea-Bissau	2012-04	Raimundo Pereira	1	1	<1
86	Mali	2012-04	Amadou Sanogo	0	1	<1
87	Central African Republic	2013-03	Francois Bozize	1	1	10
88	Egypt	2013-07	Morsi	1	1	1
89	Central African Republic	2014-01	Djotodia	1	0	1
90	Ukraine	2014-02	Yanukovych	1	0	4
91	Thailand	2014-05	Yingluck Shinawatra	0	1	3
92	Burkina Faso	2014-10	Campaore	1	1	27
93	Burkina Faso	2014-11	Traore	1	1	<1
94	Yemen	2015-01	Abd Rabbuh Mansur Hadi	1	0	3
95	Lesotho	2015-03	Thabane	1	0	3

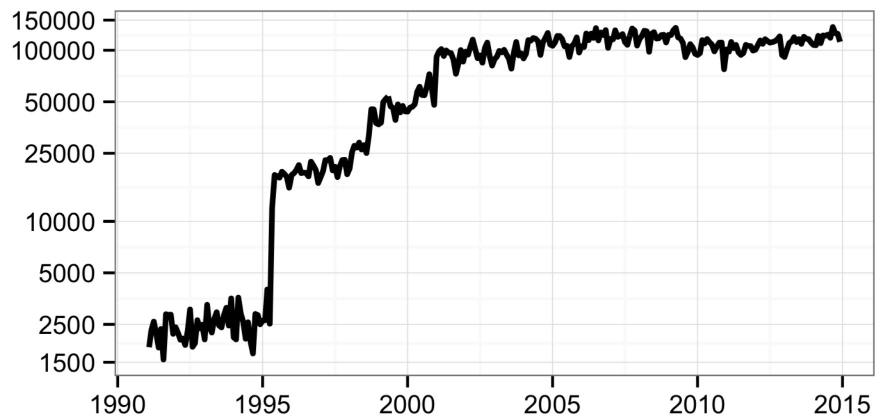

Figure A3: ICEWS events by month

3.2 Thematic models

As we mention in the main text, our contribution is not in the thematic models, and we welcome efforts to ground them more strongly in existing work. But here are brief descriptions of the models as they are, which at least should give a hint at their genesis.

All models include a variable capturing the total volume of ICEWS events in a given month, necessary as the volume of events has changed significantly from the start of the data in 1991 to the present. The monthly volumes are shown in Figure A3.

Leader Characteristics Drawing on the literature on leadership tenure (Bueno de Mesquita et al. 2005, Acemoglu and Robinson 2006, Svolik 2012), we build a model that captures the leaders' individual characteristics, as well as internal regime cooperation. The literature on leadership survival focuses on a leaders' 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, we include per capita measures of Internet users and cell subscribers into the risk equation. 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 based on the main components of the (Goldstone et al. 2010) model, which was developed to predict general instability for the PITF. The outcome, political instability, like ILCs is an aggregate of several phenomena that have largely been studies separately: civil war, adverse regime changes, mass killings, and state collapse. The so-called Global Instability model (Goldstone et al. 2010) used to predict instability is based on four variables, regime type, infant mortality, armed conflict in four or more bordering states, and state-led discrimination.

We have tried to match these four indicators as closely as possible. The first set of variables in our version are dummy indicators for specific types of regimes, derived from the Polity scheme: partial autocracy, partial democracy with factionalism, partial democracy without factionalism, and full democracy. These are coded based on the executive recruitment and competitiveness of political participation variables according to a table shown in Figure 1 in the Goldstone article. The other three variables in the original model are infant mortality, logged and normalized to global average by year, indicator for major conflict in four or more neighboring states, and state led discrimination using Minorities at Risk. We use the fraction of excluded population from EPR for state-led discrimination, and for the neighborhood risk indicator use two spatial weights of ethno-religious violence and rebellion in the nearest four neighboring countries.

Anti-regime Protests This thematic model is entirely focused on protest. Civil resistance campaigns are an effective means for achieving leadership change. The literature on both coup-proofing (Quinlivan 1999, Pilster and Böhmelt 2011) 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. Barany (2011) examines the role of militaries in countries experiencing unrest during the Arab Spring, and offers three factors that play a role in the military's decision: professionalization, the role of the military in the current regime vis other security services, and the potential impact of a successful revolt on the military's own interests. In addition to those factors that may encourage citizens to partake in mass protests, like poor governance, the military's behavior is a key determinant of a revolution's success. Revolutions thus are, like coups, explained from multiple angles, with arguments and models focusing on tipping points, mass protest, and the state's response. 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 possibility of contagion from instability in surrounding areas, which have separately been mentioned as contributing factors for coups, protests, and rebellions. It uses two spatial weights of opposition resistance and state repression in neighboring countries, weighted by centroid distance, as key indicators. The risk equation aims to capture susceptibility to contagion based on the country size and wealth, as well as the extent to which an opposition is existent and active. For example, we would expect that a small country is on average more sensitive to events in its neighboring countries than a country with a large population, in which attention is necessarily more domestically oriented, and that a wealthy country would be better able to resist cross-border influence by for example being better able to step smuggling and illegal migration that might convey conflict.

Internal Conflict The internal conflict model uses GDP per capita and the level of Autocracy 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. Powell (2012), like others (Koga 2010, Galetovic and Sanhueza 2000), suggest that if the status quo is threatened through shocks like economic crises, even the most satisfied militaries may view coups as favorable. The baseline risk is determined by GDP– wealthy countries should be better able to resist financial pressure–and population. In addition, it includes the a count of opposition resistance events as an indicator for the presence of an organized and active opposition to the current regime. For countries in the high risk set, we use the presence of inflation (via WDI), current food prices (FAO), and current oil prices as triggering indicators (using US EIA figures).

3.3 Theme model estimates

Tables A3 through A9 show estimates for the 7 thematic models.

Key for results table annotations	
^{s}	Spatial lag
^{a}	$\log_{10}(x+1)$
^{b}	$\log_{10}(x)$
^{c}	$\ln(x)$
*	Normalized

Table A3: Leader characteristics

Variable	Estimate	StdErr	p
Duration eq. intercept	-0.77	1.79	0.67
Dissident to gov't material conflict^a	-2.09	0.48	0.00
Intra-government material cooperation^a	0.04	0.88	0.96
Leader age	0.01	0.02	0.65
Global event volume^b	1.61	0.38	0.00
Hazard shape	0.47	0.11	0.00
Risk eq. intercept	-14.92	9.19	0.10
Leader entered irregularly	7.00	25.42	0.78
Leader imposed by foreign power	-15.38	27.88	0.58
Leader months in power^a	0.71	0.93	0.44
Global event volume^b	4.58	2.67	0.09

Table A4: Public discontent

Variable	Estimate	StdErr	p
Duration eq. intercept	3.20	1.47	0.03
Intra-government verbal cooperation^a	1.17	0.50	0.02
Gov't to dissident verbal conflict^a	-2.18	0.93	0.02
Dissident to gov't verbal conflict^a	-1.36	0.90	0.13
Anti-government protests^a	-0.98	0.58	0.09
Global event volume^b	0.69	0.39	0.08
Hazard shape	0.44	0.09	0.00
Risk eq. intercept	157.39	247.04	0.52
Internet users	-6.42	10.43	0.54
Mobile cellular users	-1.08	1.72	0.53
Fraction excluded from power^a	23.24	51.32	0.65
Autocracy score	1.76	2.80	0.53

Table A5: Global instability (Goldstone)

Variable	Estimate	StdErr	p
Duration eq. intercept	2.95	1.21	0.01
Ethno-rel. conflict, nearest 4 neighbors[s]	0.05	0.18	0.77
Civil war violence, nearest 4 neighbors[s]	0.00	0.03	0.96
Global event volume[b]	0.57	0.28	0.04
Hazard shape	0.44	0.10	0.00
Risk eq. intercept	-4.82	2.28	0.03
Full democracy	0.60	2.25	0.79
Other regime/in transition	1.89	2.19	0.39
Partial autocracy	18.20	4.07	0.00
Partial democracy	2.95	2.06	0.15
Partial democracy with factionalism	13.47	2.97	0.00
Fraction excluded from power[a]	59.53	27.56	0.03
Infant mortality rate[*]	8.74	3.90	0.03

Table A6: Protest

Variable	Estimate	StdErr	p
Duration eq. intercept	7.51	0.39	0.00
Ethno-religious conflict	6.99	75.09	0.93
Rebellion	-0.56	0.07	0.00
Protests, all	-0.03	0.00	0.00
Rebellion in pol. similar countries[s]	-0.06	0.02	0.02
Hazard shape	0.41	0.17	0.02
Risk eq. intercept	15.77	73.93	0.83
Domestic crisis events	-0.28	0.12	0.02
Military expenditure rate[b]	13.63	6.63	0.04
Global event volume[b]	0.27	14.72	0.99

Table A7: Contagion

Variable	Estimate	StdErr	p
Duration eq. intercept	-1.51	3.31	0.65
Distance-weighted opposition resistance[a, s]	-5.41	3.37	0.11
Distance-weighted repression[a, s]	4.15	4.01	0.30
Global event volume[b]	1.93	0.99	0.05
Hazard shape	0.30	0.11	0.01
Risk eq. intercept	6.81	2.46	0.01
Population[b]	1.96	0.93	0.04
GDP[b]	-2.60	0.74	0.00
Opposition resistance[b]	2.59	0.75	0.00
Global event volume[b]	-1.36	0.53	0.01

Table A8: Internal conflict

Variable	Estimate	StdErr	p
Duration eq. intercept	-0.73	1.68	0.66
domestic crisis × anti-gov't protest × mobile cellular users, [c]	0.32	0.10	0.00
Domestic crisis events[c]	-0.65	0.19	0.00
Anti-government protests[a]	-1.04	0.21	0.00
Mobile cellular users[c]	-0.35	0.25	0.16
Global event volume[b]	1.72	0.45	0.00
Hazard shape	0.33	0.11	0.00
Risk eq. intercept	16.93	5.16	0.00
GDP per capita[b]	-4.48	1.38	0.00
Autocracy score	0.03	0.02	0.27

Table A9: Financial instability

Variable	Estimate	StdErr	p
Duration eq. intercept	4.66	0.95	0.00
Inflation > 5%	-0.09	0.47	0.85
Food price index	0.00	0.01	0.62
Oil price	-0.13	1.93	0.94
Hazard shape	0.33	0.11	0.00
Risk eq. intercept	7.21	2.56	0.00
GDP[b]	-2.59	0.73	0.00
Population[b]	1.91	0.87	0.03
Opposition resistance[b]	2.53	0.75	0.00
Global event volume[b]	-1.45	0.54	0.01

3.4 Conventional out-of-sample test compared to test 6-month forecasts

Table A10 compares the AUC-ROC and AUC-PR statistics for conventional out-of-sample tests, where observed covariates are used to calculate out-of-sample predictions, and test forecasts in which successive 6-month forecasts are generated by iterating over the months in the test data, and without using covariates beyond the current month in the iteration. The latter matches the algorithm we use to generate the real-time forecasts.

Table A10: Comparison of monthly (m) vs. 6 month (6m) test prediction fit

Model	W	ROC m	ROC 6m	PR m	PR 6m
Ensemble		0.864	0.823	0.024	0.059
Leaders	0.22	0.651	0.741	0.021	0.054
Public Disc.	0.08	0.624	0.591	0.007	0.013
Global Instab.	0.11	0.800	0.802	0.014	0.031
Protest	0.14	0.735	0.629	0.078	0.025
Contagion	0.2	0.859	0.805	0.064	0.034
Int. Conflict	0.08	0.715	0.703	0.010	0.018
Financial	0.17	0.861	0.800	0.039	0.037

4 Data and imputation

4.1 Notes on variables

Our data consist of monthly observations for up to 173 countries over the 157 months from 2001-03 to 2014-03. We use the Gleditsch and Ward (1999) list of state system membership, but several countries drop out due to missing ICEWS event data, e.g. Malta, Serbia, Montenegro, Serbia and Montenegro, Kosovo, South Sudan, Maldives, Brunei, and Timor Leste.

Although the unit of observation is the country-month, many of the structural variables like GDP are measured annually. To use them at the monthly level we simply added them to the data by calendar year as appropriate. As a result, plotting the series for any particular country over time will produce step-like lines, where the steps correspond to changes in calendar years.

Our data include a large list of covariates, and we only list an discuss those actually used in the thematic models. The variables were drawn from several sources:

- Archigos. The Archigos data on state leaders are formatted such that each row corresponds to a leader spell in office, with exact days for entry and exit. We can thus use them with country-month units without problems. In addition to coding ILC on the basis of the Archigos irregular entry and exit variables, we also use leader age, whether the leader entered irregular or was foreign imposed, and time in power.
- World Development Indicators (WDI). Annually-measured structural variables from the World Bank, like Internet Users (`IT.NET.BBND.P2`), Cell Phone Users (`IT.CEL.SETS.P2`), Female Infant mortality (`SP.DYN.LE00.FE.IN`), GDP per capita (`NY.GDP.PCAP.KD`), Military Spending (`MS.MIL.XPND.GD.ZS`), and Population (`SP.POP.TOTL`).
- Polity data on regime types (Marshall and Jaggers 2014). The data we use are case-formatted, where each row corresponds to a specific regime in a country and records the start and end date of that regime. We convert these to monthly data, using the regime in place on the first of a month.
- Ethnic power relations (EPR). Annual. From the Ethnic Power Relations dataset (Wimmer, Cederman and Min 2009), this variable codes the share of the population in a country that is excluded from power, based on coding criteria described in the referenced paper: "[w]e categorized all politically relevant ethnic groups according to the degree of access to central state power by those who claimed to represent them" and "[f]or the present analysis, we distinguish only between power-holding groups (whatever their share of power) and the excluded population" (p. 326).
- FAO food prices. "The FAO Food Price Index is a measure of the monthly change in international prices of a basket of food commodities. It consists of the average of five commodity group price indices...", from the FAO at http://www.fao.org/worldfoodsituation/foodpricesindex/en/
- US EIA for oil prices. Daily, we aggregate to average price by month. Available from http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRTE&f=D.
- ICEWS event data. The data consist of atomic events with associated dates and locations as well as CAMEO event types–the latter are listed below. The date and location information allows us to aggregate directly to country-month observations, and we construct numerous indicators that count the number of events of a particular type between specific

actors, e.g. government and dissident. The indicators generally count event types falling into one of the 4 quad categories of material/verbal × conflict/cooperation, and further filter for specific groups of actors like the government.

- Spatial lags. The data include several spatial lags of event count indicators. These are construct using one of three methods: a weighted average of a particular event count in the nearest 4 neighbors, inverse distance weighted average of events in other countries, or a spatial average constructed using Gower distances that measure similarity on the Polity regime indicators.

4.2 ICEWS event data and quad variables

These variables are aggregations that consist of the count of ICEWS events in a country-month that meet the conditions represented by the variable name: first, the quad category for event types, and second, restrictions on the source and target actors.

The codebook for CAMEO, available from Parus Analytics, lists the codes used by CAMEO to categorize events. There are 20 codes (root-codes) for general types of events, like `MAKE A PUBLIC STATEMENT` or `PROTEST`, and each in turn contains a number of further sub-codes describing more specific types of events within the larger category, e.g. `Decline comment` or `Make positive comment`.

The 20 root codes in turn are further grouped into 4 aggregate groups, so called *quad categories*, that are defined by the combination of two dimensions: verbal versus material events on one dimension, and cooperate versus conflictual events on the other dimension. Table A11 shows the specific root codes associated with each quad category.

4.3 Imputation of missing values

We use a mix of strategies to impute missing values, based on strategies that seemed to work best based on visual examination of series with missing values: carry forward or backward extrapolation, copula imputation (Hoff 2007), and exponential smoothing state space models (Hyndman et al. 2008).

- We do not impute data for countries that are missing entire series, and rather drop such countries. We retain countries that have gaps in series, missing information at the beginning of a series, esp. for recently independent countries, or countries for which structural indicators have to be extrapolated due to delay in updating for certain sources.

- WDI structural indicators are imputed at the annual level before merging into the monthly data, using copula for gaps and state space models for extrapolation.

- For EPR and Polity, which code rates and categorical/ordinal variables respectively, we use carry-forward extrapolation of the last observed values, as copula can produce non-integer values in jumps in otherwise constant series.

- We do not impute missing event counts, and rather drop countries which have them missing. These are all countries which miss all data due to lack of coverage in ICEWS.

Table A11: CAMEO root codes and quad categories

Quad Category	Root code	
Verbal cooperation	01	Make a public statement
	02	Appeal
	03	Express intent to cooperate
	04	Consult
	05	Engage in diplomatic cooperation
Material cooperation	06	Engage in material cooperation
	07	Provide aid
	08	Yield
Verbal conflict	09	Investigate
	10	Demand
	11	Disapprove
	12	Reject
	13	Threaten
Material conflict	14	Protest
	15	Exhibit military posture
	16	Reduce relations
	17	Coerce
	18	Assault
	19	Fight
	20	Engage in unconventional mass violence

Source: CAMEO Codebook 1.1b3

4.4 Table of variable descriptions and sources

See Table A12 for a description of the variables used in the current thematic models, and their sources.

Table A12: Variable descriptions and sources

Variable (name in text)	Description	Source
Dissident to gov?t material conflict	Dissident to government material conflict events, CAMEO root codes 14 to 20	ICEWS events
Intra-government material cooperation	Intra-government material cooperation, CAMEO root codes 6 to 8	ICEWS events
Leader age	Leader age (years)	Archigos
Global event volume	Total number of global ICEWS events in a month, used to account for changes in volume over time	ICEWS events
Leader entered irregulary	Did the current leader enter power irregularly?	Archigos
Leader imposed by foreign power	Was the current leader imposed by a foreign power?	Archigos
Leader months in power	Number of months the current leader has been in power	Archigos
Intra-government verbal cooperation	Verbal cooperation between government-associated actors, CAMEO root codes 1 to 5	ICEWS events
Gov?t to dissident verbal conflict	Government to dissident verbal conflict, CAMEO root codes 9 to 13	ICEWS events
Dissident to gov?t verbal conflict	Dissident to government verbal conflict, CAMEO root codes 9 to 13	ICEWS events
Anti-government protests	Protests directed against government-associated actors, CAMEO root code 14	ICEWS events
Internet users	Internet users (per 100 people)	WDI
Mobile cellular users	Mobile cellular subscriptions (per 100 people)	WDI
Fraction excluded from power	Fraction of the population systematically excluded from access to national power	EPR
Autocracy score	Score of autocratic features of a country, 0-10	Polity
Ethno-religious conflict, nearest 4 neighbors	Coercion, assault, fighting, and mass violence involving ethnic or religious actors in nearest 4 neighbors, CAMEO root codes 17 to 20	ICEWS events
Civil war violence, nearest 4 neighbors	Coercion, assault, fighting, and mass violence involving seperatist rebel or national insurgency actors in nearest 4 neighbors, CAMEO root codes 17 to 20.	ICEWS events
Full Democracy	See Goldstone 2010 Figure 1 page 196	Polity
Other regime/in transition	Residual category for several combinations of Polity codes not captured in the Goldstone 2010 classification, mainly for transitioning, unstable, or occupied countries.	Polity
Partial autocracy	See Goldstone 2010 Figure 1 page 196	Polity
Partial democracy	See Goldstone 2010 Figure 1 page 196	Polity
Partial democracy with factionalism	See Goldstone 2010 Figure 1 page 196	Polity
Infant mortality rate	Mortality rate, under-5 (per 1,000 live births)	WDI
Low-intensity ethno-religious conflict	Protest, posturing, and reduced relations involving ethnic or religious actors, CAMEO root codes 14 to 16	ICEWS events
Low-intensity rebellion	Protest, posturing, and reduced relations between a country's government and seperatist rebels, CAMEO root codes 14 to 16	ICEWS events
Protests, all	All protests, regardless of source and target, CAMEO root code 14	ICEWS events
Rebellion in pol. similar countries	Protest, posturing, and reduced relations between a country's government and seperatist rebels in countries with similar political system, CAMEO root codes 14 to 16	ICEWS events
Domestic crisis events	Non-protest verbal and material conflict within a country's government, CAMEO root codes 10 to 20, excluding 14	ICEWS events
Military expenditure rate	Military expenditure (% of GDP)	WDI
Distance-weighted opposition resistance	Non-violent opposition–posturing, reduced relations, and protest–by seperatists, insurgents, or the opposition against the country government in nearby countries, CAMEO root codes 14 to 16	ICEWS events
Distance-weighted repression	State-led repression–coercion, assault, fighting, mass violence–against seperatists, insurgents, opposition, and opposition parties in nearby countries, CAMEO root codes 17 to 20	ICEWS events
Population	Population, total	WDI
GDP	GDP (constant 2000 US$)	WDI
Opposition resistance	State-led repression–coercion, assault, fighting, mass violence–against seperatists, insurgents, opposition, and opposition parties, CAMEO root codes 17 to 20	ICEWS events
Domestic crisis x anti-gov't protests x mobile cellular users	See the individual component variables	Multiple
Mobile cellular users	Mobile cellular subscriptions (per 100 people)	WDI
GDP per capita	GDP per capita (constant 2000 US$)	WDI
Inflation > 5%	Inflation, consumer prices (annual %)	WDI
Food price index	Index of international prices of a basket of food commodities, 2002-2004 = 100	FAO
Oil price	Europe Brent spot price (USD per barrel)	US EIA

References

Acemoglu, Daron and James A. Robinson. 2006. Persistence of power, elites and institutions. Technical report National Bureau of Economic Research.

Barany, Zoltan. 2011. "The role of the military." *Journal of Democracy* 22(4):24–35.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2005. *The Logic of Political Survival.* Cambridge, Mass: MIT Press.

Chenoweth, Erica and Maria J. Stephan. 2011. *Why civil resistance works: The strategic logic of nonviolent conflict.* Columbia University Press.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, insurgency, and civil war." *American Political Science Review* 97(1):75–90.

Galetovic, A. and R. Sanhueza. 2000. "Citizens, autocrats, and plotters: A model and new evidence on coups d'etat." *Economics & Politics* 12(2):183–204.

Gleditsch, Kristian S and Michael D Ward. 1999. "Interstate system membership: A revised list of the independent states since 1816." *International Interactions* 25(4):393–413.

Goemans, Hein E., Kristian S. Gleditsch and Giacomo Chiozza. 2009. "Archigos: A Data Set on Leaders 1875–2004." Codebook.

Goldstein, Alex, Adam Kapelner, Justin Bleich and Emil Pitkin. 2015. "Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation." *Journal of Computational and Graphical Statistics* 24(1):44–65.

Goldstone, Jack A, Robert H Bates, David L Epstein, Ted Robert Gurr, Michael B Lustik, Monty G Marshall, Jay Ulfelder and Mark Woodward. 2010. "A global model for forecasting political instability." *American Journal of Political Science* 54(1):190–208.

Hoff, Peter D. 2007. "sbgcop: Semiparametric Bayesian Gaussian copula estimation and imputation." R Package.

Hyndman, Rob, Anne B Koehler, J Keith Ord and Ralph D Snyder. 2008. *Forecasting with exponential smoothing: the state space approach.* Berlin: Springer-Verlag.

Koga, Jun. 2010. "Causes and Consequences of Coup-Proofing Strategies." Unpublished.

Marshall, Monty G and Keith Jaggers. 2014. "Codebook: Polity IV project: Political regime characteristics and transitions, 1800-2002.".

Montgomery, Jacob M., Florian Hollenbach and Michael D. Ward. 2012. ""Replication data for: Improving Predictions Using Ensemble Bayesian Model Averaging".". http://hdl.handle.net/1902.1/17286 IQSS Dataverse Network.

Muchlinski, David, David Siroky, Jingrui He and Matthew Kocher. 2016. "Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data." *Political Analysis* 24(1):87–103.

Pilster, Ulrich and Tobias Böhmelt. 2011. "Coup-proofing and military effectiveness in interstate wars, 1967–99." *Conflict Management and Peace Science* 28(4):331–350.

Powell, Jonathan. 2012. “Determinants of the Attempting and Outcome of Coups d’etat.” *Journal of Conflict Resolution* 56(6):1017–1040.

Quinlivan, J.T. 1999. “Coup-proofing: Its practice and consequences in the Middle East.” *International Security* 24(2):131–165.

Raftery, Adrian E., Tillman Gneiting, Fadoua Balabdaoui and M. Polakowski. 2005. “Using Bayesian Model Averaging to Calibrate Forecast Ensembles.” *Monthly Weather Review* 133:1155–1174.

Svolik, Milan. 2008. “Authoritarian Reversals and Democratic Consolidation.” *American Political Science Review* 102(2):153–168.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule.* New York: Cambridge University Press.

Vittinghoff, Eric and Charles E McCulloch. 2007. “Relaxing the rule of ten events per variable in logistic and Cox regression.” *American journal of epidemiology* 165(6):710–718.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. “Ethnic politics and armed conflict: A configurational analysis of a new global data set.” *American Sociological Review* 74(2):316–337.