IMF Working Paper

Forecasting Social Unrest:
A Machine Learning Approach

by Chris Redl and Sandile Hlatshwayo

INTERNATIONAL MONETARY FUND

IMF Working Paper

Strategy, Policy and Review Department

Forecasting Social Unrest: A Machine Learning Approach

Prepared by Chris Redl and Sandile Hlatshwayo [†]

Authorized for distribution by Daria Zakharova

November 2021

Abstract

We produce a social unrest risk index for 125 countries covering a period of 1996 to 2020. The risk of social unrest is based on the probability of unrest in the following year derived from a machine learning model drawing on over 340 indicators covering a wide range of macro-financial, socioeconomic, development and political variables. The prediction model correctly forecasts unrest in the following year approximately two-thirds of the time. Shapley values indicate that the key drivers of the predictions include high current and prior levels of unrest, food price inflation and mobile phone penetration, which accord with previous findings in the literature.

Contents

1   Introduction

Gideon Rachman of the *Financial Times* dubbed 2019 "the year of the street protest."[1] From Hong Kong SAR to Bolivia to Lebanon, large-scale protests broke out, disrupting daily life and economic activity. While widespread, these protests were highly varied in their alleged cause. In Chile, an increase in metro fare prices; in India, a proposed citizenship law; in Malta, claims of corruption. More broadly, the incidence of social unrest has been rising over the past decade (see figure 1a). While this upswing began with the well-known protests in Tunisia in early 2011, which spread to other Arab states, it has extended to include a wide group of countries and regions throughout the decade.[2] These include events relating to anti-austerity and Occupy movement protests in Europe in 2012, anti-government and election events in 2013 in Europe and Latin America, and election and pro-democracy protests in 2014 (Hong Kong SAR, India, Thailand). The last half of the decade saw record levels of protest events including the "Yellow Vests" protests in France, judicial reform in Poland, Catalan pro-independence in Spain and, most recently, protests against police violence in the USA.

The spectre of social unrest raises a number of policy relevant risks to financial, economic and political stability. Barrett et al. (2021) provide high frequency evidence that unrest events reduce equity valuations while Acemoglu et al. (2017) find, for the case of Egypt during the Arab Spring, equity valuation losses for firms connected to the administration currently in power - indicating that protests can play an important role in curtailing rent seeking behaviour. The impact on the real economy of unrest events can be seen in the 0.5 percentage point decline in per capita GDP around unrest events in our sample (see figure 1b). Of course, this unconditional relationship may express mere correlation or simultaneity in the link between unrest and growth. In a companion piece, Hlatshwayo and Redl (2020), we use variation of unrest in neighboring states to instrument for unrest shocks showing that these events are associated with a decline in GDP of around 1 percent over a three year period[3]. Hadzi-Vaskov et al. (2021) find that GDP declines by 0.2 percentage points following an unrest shock, using the same instrumental variable as Hlatshwayo and Redl (2020). Unrest also raises the likelihood of broader political changes. Aidt and Franck (2015) provide empirical evidence for Acemoglu and Robinson (2000)'s theory linking social unrest to the threat of revolution (and the extension of the franchise (i.e., voting rights) by elites to avoid revolution). They show that U.K. constituencies where the swing riots of 1830-31 took place were associated with a significantly larger vote share for the party that supported voting reform. Similarly Aidt and Leon (2016) demonstrate that drought-induced riots are associated with incumbents making democratic concessions in Sub-Saharan Africa. The role of in-person protest remains relevant even in advanced democracies: Madestam et al. (2013) show that the exogenous changes in attendance at "Tea Party" movement protests was associated with more Republican party votes in the 2010 midterm elections. These examples highlight a potential upside of social unrest; it may support beneficial reforms over the medium term.

These substantive macro-implications raise the question of whether we can predict unrest and the variables that drive it with a degree of reliability that is useful for policy. This paper develops a forecasting model to predict unrest events one year ahead, where unrest events are measured using the data set of Barrett et al. (2020), based on newspaper mentions of unrest. This allows us to determine what predictors are associated with future unrest and to produce an index of the risk of unrest based on the probability our model assigns to these events. A number of potential drivers of unrest have been identified in the literature: increases in food prices (Bellemare (2015)), inequality (Acemoglu and Robinson (2000)), competition between elites (Turchin and Korotayev (2020)) and population growth coupled with competition for resources in less developed nations (Acemoglu et al. (2019)). Social media penetration has been shown to aid coordination of protest activities (Enikolopov et al. (2020)) where

---

[1]Rachman, G., 2019. 2019: the year of the street protest. Financial Times, [online] Available at: https://www.ft.com/content/9f7e94c4-2563-11ea-9a4f-963f0ec7e134 [Accessed 5 November 2021]

[2]Egypt, Libya, Bahrain, Syria and Yemen.

[3]Using waves of unrest in neighboring states for identification was first suggested by Barrett et al. (2020).

that role can be amplified by economic downturns (Manacorda and Tesei (2020)) and is associated with greater regional spillovers (Arezki et al. (2020)). However, it is likely that the drivers of unrest are disparate and interact with socioeconomic conditions in complex and non-linear ways that are difficult to enumerate.

Figure 1

(a) Unrest Events

(b) Per Capita GDP growth

Source: Barret et al. (2020), WEO data base, IMF staff calculations
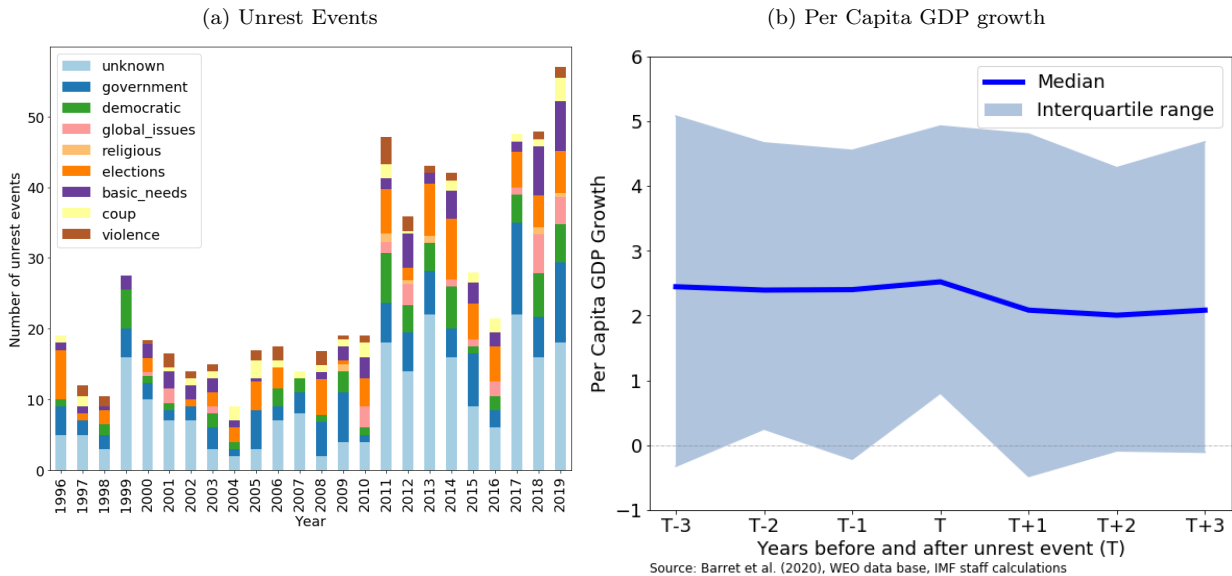
We employ a flexible machine learning approach to gauge the importance of a large set of predictors and capture non-linearities. Our preferred model has a balanced accuracy level of 66% and, in that sense, is correct in predicting unrest approximately two-thirds of the time. We find a relatively modest role for predictors in the literature, with the most important predictor of unrest in the year ahead being the level of unrest in current year. However, our results do accord with the above literature in highlighting the role of inflation (especially food inflation), unrest in neighboring states, and digital media usage (such as mobile phones) out of the large set of potential predictors considered. Our paper is the first to provide evidence of the forecasting power of a such a wide range of socioeconomic, environmental, political and macroeconomic data (over 340 indicators) and to provide evidence for a very broad set of countries (125). The closest paper to ours is that of Cadena et al. (2015), who study the ability of Twitter data to forecast newspaper-identified unrest events in Brazil, Mexico, and Venezuela but do not consider additional predictors or a broader set of countries.

The structure of the paper is as follows: section 2 outlines the data employed; section 3 describes the empirical specifications used in the forecasting horse-race; section 4 discusses the drivers of the forecasting results; section 5 discusses some country examples and distributional results; and section 6 concludes.

2 Data

2.1 Unrest events of Barrett et al. (2020)

Unrest is measured using the newspaper-based Reported Social Unrest Index (RSUI) of Barrett et al. (2020) (hereafter BANL). Their source is the Dow Jones Factiva news aggregator and covers a wide range of English language newspapers and wire services in the USA, UK and Canada. The RSUI is a monthly count of articles that include the county's name where the words "protest" or "riot" or "revolution" within 10 words of "unrest" scaled by the total number of articles in a given period.[4] BANL demonstrate a close accord between their index and two

---

[4]Additionally, articles must exclude terms likely to lead to false positives such as "vote of protest", "protest vote", "protes-

major alternative sources for unrest, the Cross-National Time-Series Data (CNTSD) database of Banks and Wilson (2020) and the Armed Conflict Location and Event Database (ACLED).[5]

Rather than forecasting the continuous RSUI index at a monthly frequency, which would pose a significant challenge in separating signals from noise as well as a narrower set of high frequency predictors with which to do it, we make predictions on whether an unrest *event* takes place in the following year. BANL produce a carefully vetted unrest event database from their RSUI index. Unrest events are defined as a local monthly peak, which is a high reading for that country's RSUI and where more than 10% of articles are on the topic of unrest.[6] Events are hand-checked resulting in 15% of events (99 of 679) failing screening, however 69% of these mis-identified events are corrected with systemic changes to the search terms for the RSUI leading to only 4.6% of events mis-identified in the final search algorithm - with the largest share of those errors relating to current flags that are in fact about past unrest. This is likely to be much less of a problem at annual frequency (which we use) as discussion in an article of events that in fact took place earlier in the year should still result in a unrest flag for that year. BANL demonstrate that their approach precisely captures well known unrest episodes such as the Arab spring of 2011, the color revolutions in formerly communist countries in the 2000s, the sequence of unrest and coups taking place in Thailand between 2006-2014 as well as the waves of unrest in Venezuela following the elections in April 2013 of Nicolas Maduro following the death of Hugo Chavez in March of that year.

Not all unrest events are alike: unrest relating to a coup d'etat may have very different economic and social implications than one relating to climate change or globalization. To add granularity we manually classify the unrest events based on short written descriptions of the events provided by Barrett et al. (2020).[7] We use seven non-exclusive categories of events based on key words. If an event description contains these key words we flag it as belonging to the relevant category - see table 1. Three categories relate broadly to the political environment: Government, Democratic-reforms, and Elections. The first captures protest directly relating to presidents, political opposition, resignations, political coalitions and impeachments. The second is broader and relates to protests around democratic reforms or rights; corruption, political reforms and the free press; and issues relating to equity topics such as gender. The third directly relates to elections, which is broken into a separate category due to the large number of events of this type. A closely-related category is protests related to basic needs capturing fiscal austerity, energy subsidies, calls for improved education and access to health care, as well as general strikes. Global issues intends to capture protests around themes that have mobilized people across many countries such as environmental, anti-war, anti-globalization and anti-immigration unrest.[8] Further categories relate to coup d'etats and assassinations; religion; and protests that involve violence. We are able to classify around two-thirds of events in this way; the remaining events are labeled as unknown. Figure (1a) shows presents the unrest events based on this classification. The most common type of unrest is related to government followed by election and democratic reform-related unrest. Global issues unrest is rare on average but has increased in frequency in the past decade.

---

tant*","anniversary", "war", "memorial" or "movie". The articles must also be tagged by Dow Jones as relating to Domestic Politics or Civil Unrest, have more than 100 words and the location tag must match the country name in the search terms.

[5]CNTSD offers a wide range of countries and a long annual time series, however it is updated with a lag relative the to BANL data (which is available with only a monthly lag), and, more importantly does not allow an analysis of the drivers of the event that text-based data does (articles around RSUI based unrest events can be classified into types which we use below). ACLED is updated at a monthly frequency but does not have comprehensive coverage going sufficiently far back in time outside of Africa - See BANL for a discussion. ACLED data can be found here: https://acleddata.com/

[6]More specifically, this requires the RSUI for a given country and month to be in the top 2% of observations for that country or the mean value is greater than four standard deviations for that country or the twenty year rolling mean is greater than four 20 year rolling standard deviations.

[7]Provided by Philip Barrett in private correspondence.

[8]We exclude the Arab Spring and its related term Tahrir Square, despite relating to multiple countries, as these protests are generally thought to be a call for democratic reform in the Middle East.

Table 1: Classifying types of unrest

Unrest type	Key words	Share of events (percent)
Government	political, anti-government, government, anti-president, president, coalition, opposition, resignation, resigns, impeachment	17.5
Democratic-reform related	Arab Spring, journalist, journalists, freedom, lawyer, democracy, Tahir Square, law, independence, anti-police, constitution, anti-corruption, corruption, reform, anti-segregation, constitutional, suffrage, women, referendum, fraud, civil society	9.4
Global issues	occupy, anti-WEF, anti-Davos, anti-U.N., anti-US, intervention, foreign, anti-globalization, G20, climate, environment, environmental, immigration, Brexit migration, migrant, refugee, human rights, summit, anti-war	3.6
Religious	anti-blasphemy, Mosque, Quran	0.6
Elections	candidates, vote, electoral, poll	14.0
Basic needs	anti-austerity, austerity, electricity, energy, yellow vests, gas, strike, union, healthcare, education, school, land, agriculture	8.3
Coup/Sudden End to Tenure	ousted, assassination, assassinated, military	5.7
Violence	deadly, riots, violent, civil war, burning	4.9
Unknown	...	36.0

2.2 Predictors

A wide range of socioeconomic, environmental, political and macroeconomic data are used as inputs to forecast unrest events. Altogether, we consider over 340 indicators and a complete list is provided in the appendix; here we outline the data sources and broad types of variables covered.

Our starting point is the fiscal crisis prediction model database of Hellwig (2020), which draws primarily on the IMF's World Economic Outlook (WEO) and International Financial Statistics (IFS) and World Bank World Development Indicators (WDI) datasets. This includes a number of categorical variables splitting the sample of countries by levels of development (Advanced, Emerging and Low Income Economies), region, continent, membership in a monetary union, commodity exporting status, fuel importer classification, whether an election has been held in a given year, or if a country is part of the Heavily Indebted Poor Country (HIPC) initiative. Macroeconomic variables cover GDP level and growth, exchange rates, terms of trade, foreign reserve levels, fiscal balance, aid receipts, remittances, and population. Transformations of these variables include growth rates, first differences, lags, and volatility measures (e.g., of exchange rates by using a higher frequency version of the variable and capturing its fluctuations in a given year).

We augment this with data from the International Country Risk Guide and the Cross-National Time-Series Data Archive (CNTSD) to cover political characteristics of countries. Recent literature highlights how low-quality institutions (Moseley (2015)) can catalyze unrest while certain political regimes help mitigate protest participation (Ackermann (2017)). Given the difficulty of capturing institutional quality with just one measure, we include a battery of indicators (e.g., legislative effectiveness, executive effectiveness, government cohesion, perceptions of legitimacy and accountability, and bureaucratic quality). Regime type is measured using Polity IV scores, differences in executive selection processes, the degree of parliamentary responsibility, the nature of the head of state, and how active military groups are in political processes). These datasets also contain indicators that proxy for sources of tension (e.g., religious frictions, ethnic conflicts, war, and terrorism); polarization (e.g., the extent of political party fractionalization and the occurrence of government purges); and instability (e.g., assassinations, political violence, and mass strikes)—all of which have the potential to go hand-in-hand with broader societal unrest. Finally, we

include measures that capture the size of prison populations, which may serve as a proxy for political discord (e.g., where there are large numbers of political prisoners), an excessively strict legal system that could foment eventual uprisings, or inadequate investments in education and community development (i.e., where insufficient economic opportunities drive up both crime and unrest). With respect to the latter, indicators that measure educational attainment, poverty, quality of healthcare, and the investment environment are also included.

These data sets also allow us to proxy for informational frictions and coordination costs. Access to news media, social media, and associated technologies (e.g., televisions and mobile phones) can improve the flow and aggregation of information and ease logistical hurdles for mobilizing large populations (Manacorda and Tesei (2020)). To capture these features, we include measures of internet and mobile phone penetration as well as televisions per capita and newspaper circulation.

From the World Bank World Development Indicators (WDI) we collect data on five broad categories. First, poverty, measured via the poverty gap at the $5.5 per day level, prevalence of food insecurity and multidimensional poverty head count and intensity indices. Second, access to basic services such as electricity, sanitation and health care. Third, inequality via Gini indices, share of income held by highest 10%, financial access of the poorest 40%, as well as measures of gender inequality relating to schooling and literacy. Fourth, measures of population urbanization and density since densely populated urban environments are more likely support the spread of unrest. We also include the share of international migrants in the population to proxy for xenophobic tensions. Finally, unemployment for the population as well as the youth unemployment rate by gender are included.

We include natural disaster events from the EM-DAT database from the Centre for Research on the Epidemi-ology of Disasters at the University Catholique de Louvain.[9] Disasters, if perceived to be poorly prepared for or responded to, may lead to anti-government sentiment. We include data on the number of people affected by drought, earthquake, epidemic, extreme temperatures, floods, industrial accidents, and storms.

Elevated levels of policy and economic uncertainty have been shown to drag on growth, investment and raise unemployment (see, for example, Bloom (2014)). Besides the direct economic effects of uncertainty, the inability to plan may inhibit households and businesses and lead to tensions with the government. We include a text-based measure of uncertainty based on the Economist Intelligence Unit's quarterly country reports produced by Ahir et al. (2018) - this data set covers more than 140 countries from 1996 onwards.

Sharp increases in the cost of living are likely to lead to hardship and discontent, especially for the case of necessities such as food (see Bellemare (2015)). To capture the role of different categories of prices we use the annual maximum of the year-on-year monthly inflation figures for a variety of CPI categories from the IMF CPI database. [10] Similarly, announced policy changes, such as privatization of state-owned-enterprises or relaxation of labor market regulations may elicit protest from affected groups. To account for this channel, we include the data of Duval et al. (2018) who build narrative indicators of structural reforms covering product market and labor markets.[11]

3 Model

3.1 Machine Learning Models Considered

We consider a variety of forecasting models to predict unrest events, falling into three groups. All models are

---

[9]https://emdat.be/

[10]Available at https://www.imf.org/en/Data

[11]The authors describe the reforms as, "reducing administrative burdens and barriers to entry in product markets, including in the areas of retail trade, professional services, and some network industries; easing hiring and dismissal regulations for regular workers; increasing the ability of and incentives for the non-employed to find jobs by reducing the level or duration of unemployment benefits, or by increasing the resources for, and the efficiency of, active labor market policies; cutting labor tax wedges; and targeted policies to boost the employment rate of underrepresented groups in the labor market, including youth, women, older workers and, in some countries, migrants." See Duval et al. (2018), p.4.

described formally in the appendix; here we provide intuition for the models considered. The first group, familiar to economists, are linear regression models that use a logistic function to transform the linear model's output to ensure that the range of the function lies between 0 and 1. However, we include a regularization term that allows many predictors to be considered without suffering from over-fitting. The Ridge model adds a quadratic penalty term (L^2 norm) in the parameters of the linear regression to the standard likelihood function while Lasso adds an absolute value penalty (L^1 norm). The Ridge model assumes that the true model is dense, where all predictors matter, whereas Lasso assumes a sparse model, where only a few predictors are important.

Other models are adept at capturing non-linearities between the target and the predictors. This group includes a neural net and a support vector machine (SVM) – both are models that have performed well on non-traditional data such as text and image recognition tasks in other applications (For instance, see Gentzkow et al. (2019), Redmon and Farhadi (2017)). SVMs classify events using a non-linear decision boundary that aims to maximize the gap between different groups of observations.[12] Neural nets are closely related to standard logistic regressions except that neural networks use the output of one logistic model as an input to other logistic models, layering many on top of on another. This layering results in the ability to capture highly non-linear functions in a computationally efficient way.

The final group of models are tree based. These methods start with a decision tree - which partitions the space of predictor variables so that being in a given region has predictive power for the target variable.[13] While intuitive, such models are prone to over-fitting and so a variety of techniques have been introduced to combine a large number of trees in a robust way. AdaBoost and Gradient Boost are boosting algorithms which are algorithms that build strong prediction models from a group of weak, typically very simple, prediction models. Weak prediction models are attractive since they are not prone to over-fitting, however their individual performance is poor and so they are combined to yield more accurate predictions. Freund and Schapire (1997) introduced the AdaBoost algorithm[14] which sequentially evaluates simple models (in this case shallow tree models) placing higher weight on models that perform well while also putting higher weight on the mis-classified data. The final prediction is a weighted average of all the simple models. Gradient boosting algorithms sequentially improve simple models by also modeling the slope of the loss function found in the previous step (intuitively, this is like building a simple model to predict a target and a model to predict the errors of that model before combining them). For a popular approach to gradient boosting, see Chen and Guestrin (2016). An alternative to boosting is bagging or bootstrapping, where many models are fitted on bootstrapped samples of the training data and average the results.[15] Bagging reduces over-fitting in flexible models such as decision trees, reducing the variance of (out-of-sample) predictions. The major innovation in this area is that of Breiman (2001), who averages the performance of many de-correlated trees - where each tree is grown using only a random subsample of the available predictors (See Algorithm 1). We find that tree-based models perform best in our context with Random Forest providing the best overall model; although the Random Forest approach is comparable to the boosting methods (see 2a) and discussion below.

3.2 Model Evaluation and Performance

We evaluate the models using an expanding window split of the training and testing set starting from an initial training set of annual data from 1990 to 1995 to make a prediction for the test set of 1996. We then roll the

---

[12]The simplest case being linearly separable data where a straight line can be used for the decision boundary. In this case, an SVM would choose the parameters of a linear regression line so as to maximize the area around the regression line where no target variables overlap (i.e. maximize the gap between the two classes). As in linear regression, transformations of the original data (using basis functions) capture non-linearities.

[13]An simplified example of a tree would be if food price inflation is above a cut-off level, say 10%, then the model predicts that unrest would occur. Similarly, the tree would partition all the other predictors mentioned above depending on how large a tree a researcher has decided to build.

[14]Useful explainer from the authors is here: https://rob.schapire.net/papers/explaining-adaboost.pdf and https://mccormickml.com/2013/12/13/adaboost-tutorial/

[15]In the case of decision trees with a discrete outcome, averaging predictions takes the form of a majority vote or averaging the probability of an event from each tree. We use the toolkit of Pedregosa et al. (2011) which uses the latter approach.

sample forward, generating predictions up to the year 2019. Hyperparameters are chosen using a coarse grid using the average model performance over the full rolling window sample, i.e. up to 2019, see appendix I. We do not pursue K-fold cross validation, as is common in the machine learning literature, due to time dependence in the data which is typical in economic datasets. Additionally, BANL demonstrate significant persistence in the RSUI measure, which is the basis for the events we aim to predict. Shao (1993) demonstrates asymptotic inconsistency for K-fold cross-validation in the presence of time dependence. Creative solutions to using cross-validation in a time series context exist, such as Racine (2000) who uses a gap between the testing and training set to break serial dependence. However, this approach would consume too many degrees of freedom in our relatively short annual data set.[16]

Algorithm 1 Random Forest Classification

For each tree, T_b, for $b = 1$ to B:

1. Draw a bootstrap sample Z from the N predictors in the training sample, $X' = (X_1, X_2, ..., X_N)$.

2. Grow a random-forest tree T_b on the sample Z: recursively repeat the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached:

   (a) Select m variables at random from the p variables available

   (b) Pick the best variable and split point from among the m variables using the CART algorithm of Breiman et al. (1984):

      i. Let the data at node m be partitioned into $Q_m^{left}(\theta)$ and $Q_m^{right}(\theta)$ where $\theta = (j, s)$ is a candidate split threshold (s) for feature (j). The split is computed from the loss function: $\theta^* = \arg\min_\theta \left[\sum_{k=left,right} n_m^k p_m^k (1 - p_m^k) \right]$, where n_m^k is the share of observations in region k, p_m^k is the share of unrest events in the region $Q_m^k(\theta)$ and $k = left, right$

      ii. Repeat for subsets $Q_m^{left}(\theta^*)$ and $Q_m^{right}(\theta^*)$ until a single observation remains at node m.

   (c) Split the node into two daughter nodes.

The output of the above is a collection of B trees, $\{T_b\}_1^B$. The probability for a new data point, x, resulting in an unrest event, is given by $\hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{b=1}^B \hat{f}_{T_b}(x)$, where $\hat{f}_{T_b}(x)$ is the share of unrest events for the terminal node associated with x for the bth tree.

Our preferred test statistic is the area under the Receiver Operator Characteristic (ROC) curve (which we denote by AUC). The ROC plots the true positive rate (events correctly classified as unrest divided by total unrest events) against the false positive rate (events incorrectly classified as unrest divided by all events with no unrest associated with them) as the threshold for classifying an event as unrest (e.g., the estimated model probability is above 50%) is varied. Our benchmark is a random guess (coin flip), which is reflected by the AUC's 45 degree line (see figure 2b). The AUC is defined following the empirical AUC, denoted $\hat{\theta}$, of DeLong et al. (1988)[17]:

$$\hat{\theta} = \frac{1}{NM} \sum_{j=1}^N \sum_{i=1}^M H(X_i, Y_j)$$

---

[16]To use Racine (2000)'s methods we would need to leave a large enough gap between the training and test set to be confident we have no time dependence remaining in the data. In our case, this could be as high as five years or more given some of the data in our sample, such as the fiscal variables.

[17]The population AUC is given by $E(\hat{\theta}) = \theta = \Pr(Y < X) + 0.5\Pr(Y = X) + 0\Pr(Y > X)$. Below we will generate sample variation in our empirical AUC using the out-of-sample testing sets.

$$\text{where } H(X,Y) = \begin{cases} 1 & Y < X \\ 0.5 & Y = X \\ 0 & Y > X \end{cases}$$

Where X denotes the model probability of unrest for events where unrest took place and Y, the model probability of unrest (risk score) for an event where unrest did not take place. The intuition behind this definition is that we would want the model to assign a lower probability to events where unrest did not happen than to those where it did, and so for these cases its score rises by $1/MN$; in cases where these are equal, the model receives a lower score of $1/2MN$; and in the erroneous case (Y>X), the model receives no addition to its AUC. As noted by DeLong et al. (1988), this definition of the AUC is equivalent to the Mann-Whitney U test which estimates the probability, $\hat{\theta}$, that a randomly selected event from the group where no unrest took place will have a risk score less than or equal to a randomly selected event from the group where unrest did occur. Additionally, the AUC can, in our case with only two classes, be interpreted as a measure of balanced accuracy or the arithmetic average of the true positive rate and the true negative rate.[18] DeLong et al. (1988) provides a test statistic for comparing two models via $\hat{\theta}$, which turns out to be a simple z-score computed as $\frac{\hat{\theta}_A - \hat{\theta}_B}{\sqrt{\mathrm{Var}\left[\hat{\theta}_A - \hat{\theta}_B\right]}}$. Below we use this to test each model against the AUC equivalent of a random guess, 0.5; thus, we use the z-score that corresponds to $\frac{\hat{\theta}_i - 0.5}{\sigma_i}$, where i corresponds to the model type and σ is the standard deviation of $\hat{\theta}$ over the test set.

The model results are compared in figure 2a which shows the mean $\hat{\theta}$ along with the standard deviation over the test set (the 23 out-of-sample years, 1996-2019). Details on hyper parameter tuning are found in the appendix. All models are adept at incorporating our large set of predictors, as noted above; however, the linear models perform relatively poorly with an AUC of 0.55 for both the Lasso and Ridge logistic models. The neural net model performs comparably but remains close to a random guess (i.e., an AUC of 0.5) at 0.54. While neural nets have shown remarkable performance in contexts with very large datasets, it is somewhat unsurprising that they perform less well in our context which is too small to take advantage of the flexibility of this class of models. The SVM model preforms significantly better than the preceding models. The SVM from our hyper parameter grid-search uses a radial basis function kernel[19] to capture non-linearities; however, the favored parameterization acts to push the model towards a more linear representation reducing over-fitting - this balancing act works relatively well compared to the case of the neural net. The tree based models outperform our other models and all have (one-sided) DeLong test statistics that are significantly different from 0.5, with Random Forest possessing the highest mean $\hat{\theta}$, and is our preferred forecasting model used for the remaining results.

---

[18] Accuracy, or the share of correctly classified observations, is equivalent to balanced accuracy in the case where there are equal numbers of positive (unrest) and negative (no unrest) observations. However, in the case of unrest events being rare, as they are in our sample, balanced accuracy is a preferred measure since accuracy can be very high in such a case with a simple but uninformative rule such as always assume a crisis does not occur. Specifically, Accuracy $= \frac{TP+TN}{TP+FN+TN+FP}$ where as Balanced Accuracy$= \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$, where TP is true positives, FP is false positives, TN is true negatives and FN is false negatives. Balanced accuracy is also related to the signal-to-noise ratio used in the economics literature, e.g. Kaminsky and Reinhart (1999), to find threshold values for predictors of a crisis (e.g. banking or currency) by maximising the ratio.

[19] Defined between two observations in the set of predictors, this kernel is $K(x, x') = \exp(-\gamma \parallel x - x' \parallel^2)$

Figure 2
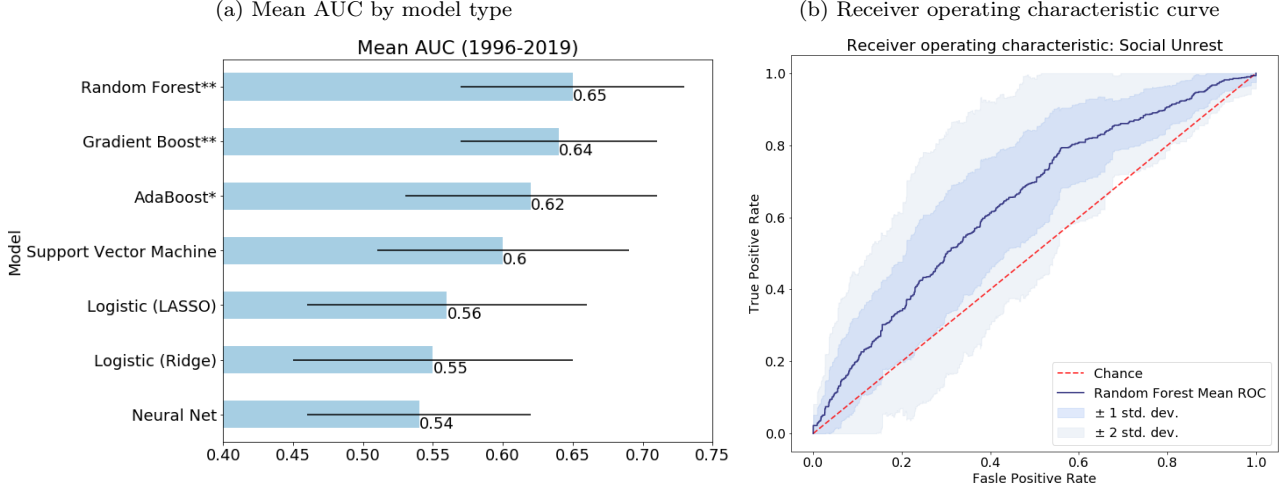
(a) Mean AUC by model type

(b) Receiver operating characteristic curve

**significant at 5%, * at 10% for DeLong test. Horizontal bars show mean AUC over the test set with the error bars representing standard errors.

The broad picture of our preferred model's performance is demonstrated by the ROC curve (see figure 2b). This indicates that while our model performs significantly better than a coin-flip, the trade-off involved requires accepting a relatively high false positive rate (40-60%) to achieve true positive rates in the 60-80% region. To make this more concrete we calculate an optimal threshold at which to judge whether unrest has occurred. We pick the threshold to minimize the average of missed unrest events (1- true positive rate) and false alarms, for each out-of-sample prediction. The average probability of an unrest event in a country in a year is 16.1%, and the average threshold selected from this procedure is slightly higher at 16.8%. The choice of a threshold allows us to compute some intuitive model performance statistics. Balanced accuracy is 65.9%, which along with the result for the mean AUC above indicates that this model is correct approximately two thirds of the time. Recall measures the number of relevant events selected and is defined as $\frac{\text{TP}}{\text{TP+FN}} = 70.9\%$; thus, few crises are missed. However, this is at a cost of relatively low precision, defined as $\frac{\text{TP}}{\text{TP+FP}} = 30.8\%$ indicating that a significant number of false alarms in order to reach that level of recall. This may seem odd given the way the threshold is chosen but it is simply an expression of the trade-off illustrated in the ROC curve: a relatively high false positive rate is required to catch a large number of unrest events.

The performance of the model varies over the sample period, with highest performance in the early 2000s and late 2010s (see figure 3a). It is noteworthy that the worst performance takes place around 2009-10 when the model is surprised by the structural break induced by the Arab spring; however, it recovers in the following years. We also calculate the model's performance by type of unrest event where the model performs similarly on most categories with the exception of violent unrest and coups (see figure 3b). Hlatshwayo and Redl (2020) show that coups and violent unrest tend to be associated with larger negative macroeconomic consequences and, based on our classifications of unrest types, violent unrest is less prevalent than other forms (e.g., anti-government and election-related unrest). This work suggests that the larger economic effects from violent forms of unrest may be associated with their rare, unanticipated, and unpredictable nature.

Figure 3

(a) Area under receiver operating characteristic curve by year
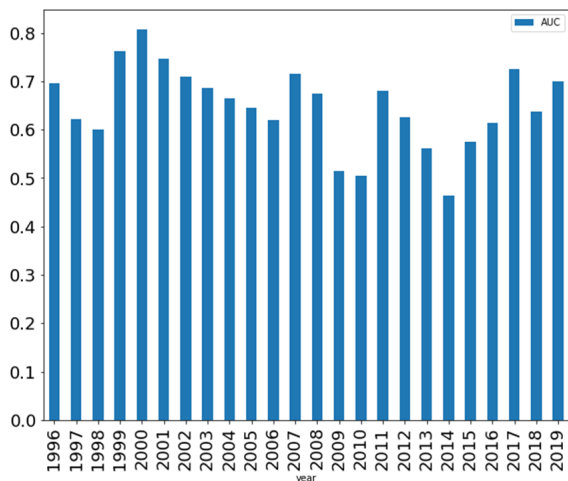

(b) Mean balanced accuracy by type of unrest
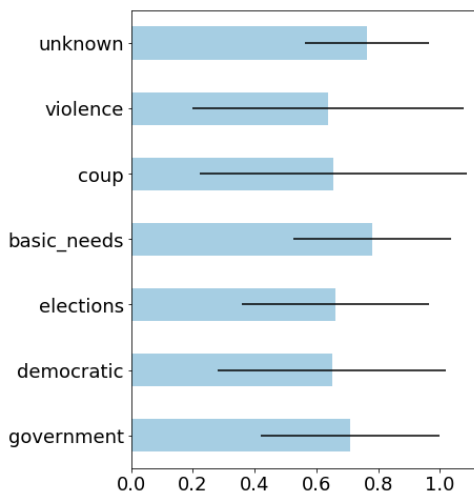
**significant at 5%, * at 10% for DeLong test. Horizontal bars show mean AUC over the test set with the error bars representing standard errors.

4 Feature Importance

To understand the importance of the different features in driving the model predictions, we use Shapley values. The description of Shapley values follows Joseph (2019). To understand this approach to feature importance in non-linear models, it is useful to begin with the linear case. The contribution of a feature, or a single observation of that feature, to a prediction in a linear model is given by the regression coefficients multiplied by the observation value(s). Consider the linear model; the model decomposition Φ for the feature x, with observation, i, is:

$$\Phi\left[\hat{f}(x_i)\right] = \phi_0(\hat{f}) + \sum_{k=1}^{n} \phi_k(x_i; \hat{f}) = \hat{\beta}_0 + \sum_{k=1}^{n} x_{ik}\hat{\beta}_k \tag{1}$$

where $\hat{\beta}_0$ is the estimate of the unconditional expected value of $f(x)$, and β_k the slope coefficients corresponding to the kth feature. Shapley values provide a generalization of this formula for non-linear models drawing on ideas from cooperative game theory. Each observation is treated like a player in a cooperative game with other observations where the result of the game is a model prediction. The difficulty is that pay-offs to observations working together cannot be assessed by simply using a single observation. This is because non-linear models, usually rely on interaction effects between many observations. Adding an observation that is highly correlated with the observations already in a coalition will lead to little additional pay-off but uncorrelated observations that contain useful signals will. The way to gauge the value of these coalitions is formalized below:

$$\hat{f}(x_i) = \phi_0^S + \sum_{k=1}^{n} \phi_k^S(x_i; \hat{f}) \equiv \Phi^S(x_i) \tag{2}$$

$$\phi_k^S(x_i; \hat{f}) = \sum_{x' \subseteq \{x_1, x_2, ..., x_p\}\setminus\{x_k\}} \frac{|x'|!(n - |x'| - 1)!}{n!}\left(\hat{f}(x_i|x' \cup \{x_k\}) - \hat{f}(x_i|x')\right) \tag{3}$$

where $x' \subseteq \{x_1, x_2, ..., x_p\}\setminus\{x_k\}$ is the set of all variable coalitions when we exclude the kth variable (whose

value we wish to measure), $|x'|$ is the number of variables included in the model evaluation for each set of coalitions, $\frac{|x'|!(n-|x'|-1)!}{n!}$ is the combination weighting factor for each coalition and $\hat{f}(x_i|x' \cup \{x_k\}) - \hat{f}(x_i|x')$ is the pay-off for including x_k in the coalition x'. This approach to measuring feature importance in the machine learning literature is due to Strumbelj and Kononenko (2010).

Figure 4: Shapley values

(a) Average Shapley value contribution, 1996-2019

(b) Time variation in selected Shapley values
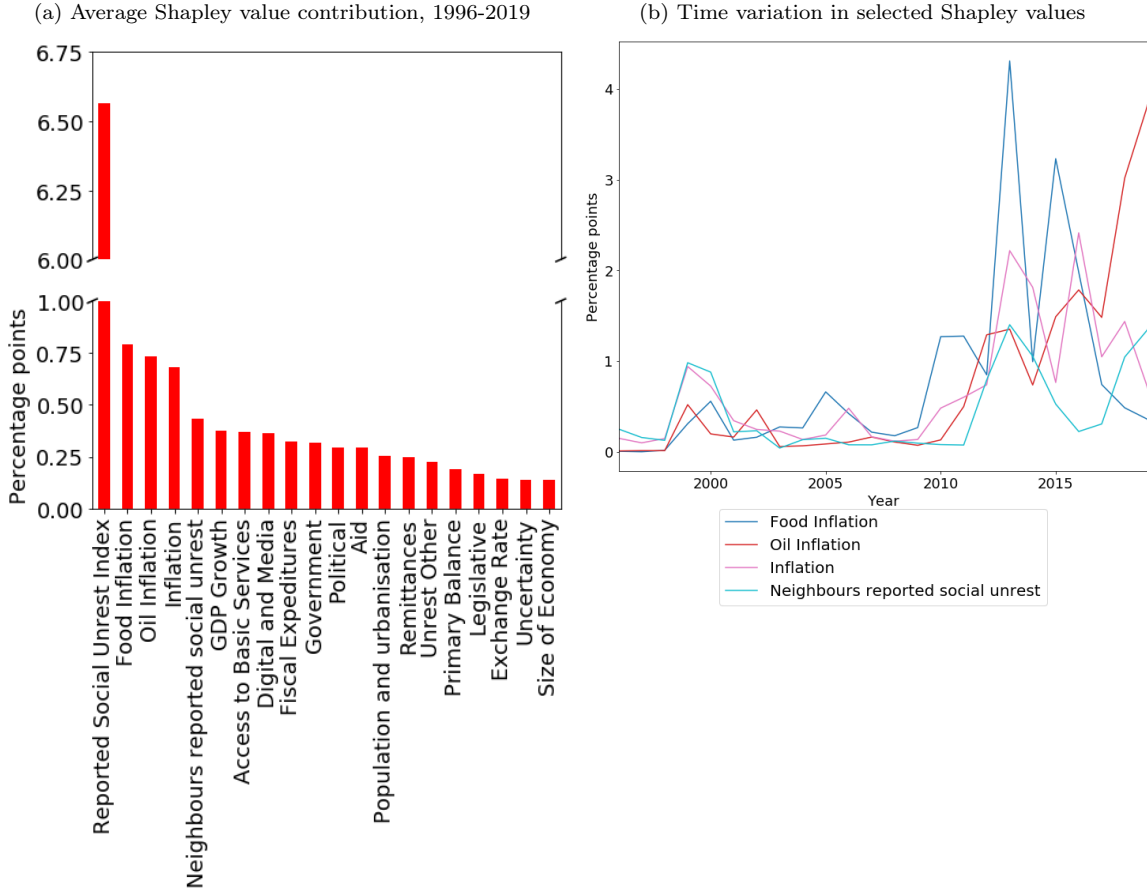
We calculate Shapley values using the implementation of Lundberg and Lee (2017). A choice needs to be made as to whether the model is evaluated on the data on which it is trained or on the out of sample data on which its forecasts are evaluated (the test data). Here we apply the decomposition (3) in the same way we evaluated the model's forecast performance: fitting the model on an initial sample of 1990-1995 and calculating the Shapley values corresponding to the predictions for 1996, then rolling the sample forward, one year at a time, to 2019. An alternative approach would be to use the model fitted on the full sample of data, i.e., up to 2019. We choose the test sample approach since highly flexible machine learning models such as Random Forest can have arbitrarily good in-sample fit, which is unlikely to be representative of out-of-sample predictions. This approach is also appealing in that it provides a sense of how the importance of different features has changed over time as the model is re-estimated, as it would be done in practice.

Due to the large number of features, we aggregated them into categories to gauge the importance of different categories of variables, the mapping is available in the appendix.[20] For example, the Reported Social Unrest Index of Barrett et al. (2020) and all its lags are pooled into one category of "Reported Social Unrest Index" (RSUI) which captures the influence of past levels of unrest on future unrest; the features covering internet, television, newspaper,

---

[20]One of the appealing properties of Shapley values is that they are additive; see Joseph (2019).

telephone and mobile phone usage are grouped under "Digital and Media". Figure (4a) presents the average of the absolute values of the Shapley values over all test sets by category.[21] The values represent the contribution to predictions over and above the average prediction $\left(\phi_0^S\right)$, which is 16 percent for all years and countries. The RSUI is by far the most important category of features adding, on average, around 6.5 percentage points to the average prediction. The remaining categories yield smaller contributions. The next most important group are those relating to inflation, where food and oil inflation are broken out separately given their importance. The influence of food inflation is in line with the findings of Bellemare (2015). Oil price inflation is primarily influential in its interaction with different levels of the RSUI, where lower values of oil inflation dampen the effect of high historical levels of unrest. We also find a role for contagion in driving unrest where unrest in neighboring states can help predict unrest, in line with Arezki et al. (2020) and Barrett et al. (2020). The importance of Digital and Media is in line with the findings of Enikolopov et al. (2020) and Manacorda and Tesei (2020).

Figure (4b) illustrates that the influence of features can change abruptly over time, with a large role for inflation (especially food price inflation) during the sharp increase in unrest following the Arab Spring of 2011. The model also indicates the contagion from neighbors became more important around that time as well. Inflation has remained important in the 2010s; however, the role of food prices has declined while that of oil prices has increased, a pattern broadly in line with their respective price changes over this period.

5   A Social Unrest Risk Index

The model produces a probability of unrest for each country for each year. This probability can be treated like a risk index - an indicator of the likelihood of unrest in the following year. The average risk for the sample is 16 percent but this masks a pattern of generally lower risk in the first half of the sample, hovering around 12 percent, then rising significantly after the Arab Spring and maintaining elevated levels of the RSUI in the 2010s of values closer to 20 percent (Figure 5a). This shift towards a higher risk of unrest can be seen in the shifts of the distribution of unrest risk over all countries in the sample. Figure (5b) shows kernel density plots of the distribution of unrest risk over time. In the first half of the sample, the majority of countries had a low risk of unrest with a long right tail of higher risk countries. By 2010, risks had risen in the tail and a bi-modal profile emerged. After 2015, the distribution remains bi-modal but the mode moved to the right of the mean and we see a much larger incidence of high risk countries than in the 1990s. In summary, unrest moved from a low probability tail risk in the 1990s (of around 10 percent on average) to a relatively likely outcome by the end of the 2010s (at 25 percent).

---

[21]Shapley values can be both positive and negative; here we use the absolute value to gauge influence in both directions.

Figure 5: Risk of Unrest

(a) Average probability of unrest   (b) Distribution of unrest risk over time
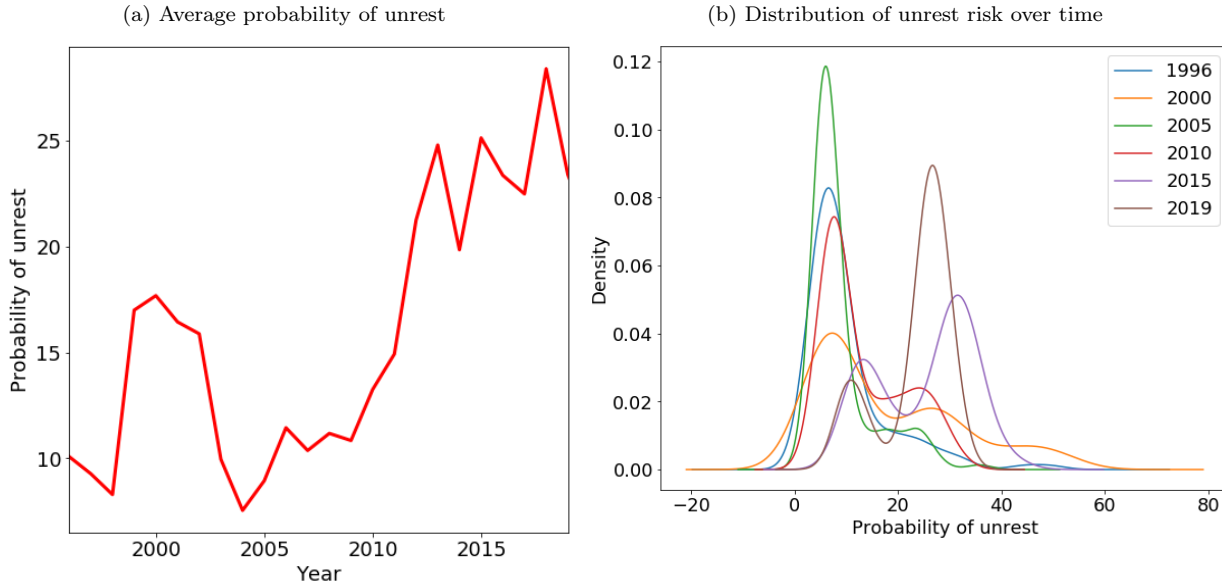
(a) Rolling out-of-sample probability of unrest averaged across countries for each year. (b) Density plot for the probability of unrest across countries.

Figure (6) illustrates the risk index for selected country cases and contrasts this with unrest events. Unrest events flagged in the figure correspond to one year *before* they occur; this is done because the model is producing a risk index based on the probability unrest will occur in the following year. Therefore changes in the risk index are then aligned with the dates when unrest takes place. The event timing discussed here will correspond to the figure. Both the United Kingdom and the United States share a similar pattern with a low risk prior to the last decade when risks rose throughout the decade. The United Kingdom experienced a large number of unrest events: the global events relate to the Occupy movement and Brexit; democratic events relate to the Scottish independence referendum; and basic needs unrest relate to anti-austerity protests. The risk index did not provide warning of the global events relating to the Occupy protests but did remain elevated from 2012 foreshadowing the frequent protests taking place thereafter. The labeled events in the United States relate to the election of President Donald Trump and protests against police violence (under Democratic unrest). Notably, the risk index rose prior to these events in 2015 and remained high thereafter. Egypt and Thailand offer cases where protest led to a change of government. In the case of Egypt, we have the Tahrir Square protests and subsequent resignation of President Hosni Mubarak. The model did not pick up the risk of this event, as the model does poorly in catching the structural break that takes place around the time of the Arab Spring in general. Part of the explanation may be that the model underappreciated the importance of food prices prior to the Arab Spring but subsequently places a high weight on this feature. The model performs better around the events leading up to the coup d'etat in Thailand that took place in 2013, with a high and rising risk index (with the exclusion of 2012).

14

Figure 6: Country Cases

(a) United Kingdom

(b) United States

(c) Egypt

(d) Thailand

6   Conclusion

Social unrest raises financial, economic and political risks in the near term, with potentially positive effects over the medium term. The increasing frequency of such events over the past decades suggests that an early warning system for these events may be useful for policy makers. While a few indicators have been identified in the literature as important in driving unrest, our work illustrates that the drivers of unrest are disparate and interact with a wide range socioeconomic conditions in complex and non-linear ways. In particular, we combine the measurement of unrest by Barrett et al. (2020) with a large data set of socioeconomic, environmental, political and macroeconomic data to forecast unrest events. We compare the performance of a set of popular forecasting models, finding that the linear models typically used by economists perform poorly as do highly complex models popular in Artificial Intelligence (due to our relatively small data set). Tree based models appear offer the appropriate balance of flexibility and simplicity. Our preferred Random Forest model produces a balanced accuracy level of 66%[22] and an AUC of 65%, which is significantly better than chance, based on the test of DeLong et al. (1988).

We explore the drivers of the model's predictions using Shapley values. These accord with prior literature in emphasizing food price inflation, media usage and contagion from unrest in neighboring states. However, this paper also puts in context a wider variety of drivers such as GDP growth, development indicators such as access to basic

---

[22]Conditional on our threshold set to minimize the average of missed unrest events and false alarms.

services, urbanization, remittances, and uncertainty. We find that there is a substantial auto-regressive nature to unrest where unrest levels today (and in the past) are the most important predictor for future unrest.

The predictions of the model provide an unrest risk index that shows a rising risk of unrest since the early 1990s when the risk of unrest was around 10 percent to a level closer to 25 percent today. Future work could focus on using these results to build higher frequency predictions. Our results suggest that incorporating social media data and high frequency data on inflation may be fruitful.

References

Acemoglu, D., Fergusson, L., and Johnson, S. (2019). Population and Conflict. *The Review of Economic Studies*, 87(4):1565–1604.

Acemoglu, D., Hassan, T. A., and Tahoun, A. (2017). The Power of the Street: Evidence from Egypts Arab Spring. *The Review of Financial Studies*, 31(1):1–42.

Acemoglu, D. and Robinson, J. A. (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *The Quarterly Journal of Economics*, 115(4):1167–1199.

Ackermann, K. (2017). Individual differences and political contexts the role of personality traits and direct democracy in explaining political protest. *Swiss Political Science Review*, 23:21–49.

Ahir, H., Bloom, N., and Furceri, D. (2018). World uncertainty index. *mimeo*.

Aidt, T. S. and Franck, R. (2015). Democratization under the threat of revolution: Evidence from the great reform act of 1832. *Econometrica*, 83(2):505–547.

Aidt, T. S. and Leon, G. (2016). The democratic window of opportunity: Evidence from riots in sub-saharan africa. *Journal of Conflict Resolution*, 60(4):694–717.

Arezki, R., Adessé Dama, A., Djankov, S., and Nguyen, H. (2020). Contagious protests. *World Bank Policy Research Working Paper 9321*.

Banks, A. and Wilson, K. (2020). Cross-national time-series data archive.

Barrett, P., Appendino, M., Nguyen, K., and de Leon Miranda, J. (2020). Measuring social unrest using media reports. *IMF Working Paper No. 20/129*.

Barrett, P., Bondar, M., Chen, S., Chivakul, M., and Igan, D. (2021). Social unrest and financial markets. *forthcomming IMF Working Paper*.

Bellemare, M. F. (2015). Rising food prices, food price volatility, and social unrest. *American Journal of Agricultural Economics*, 97(1):1–21.

Bloom, N. (2014). Fluctuations in Uncertainty. *Journal of Economic Perspectives*, 28(2):153–176.

Breiman, L. (2001). Random forests. *Machine Learning*.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984). *Classification And Regression Trees*. CRC Press.

Cadena, J., Korkmaz, G., Kuhlman, C. J., Marathe, A., Ramakrishnan, N., and Vullikanti, A. (2015). Forecasting social unrest using activity cascades. *PLOS ONE*, 10(6):1–27.

Chen, T. and Guestrin, C. (2016). Xgboost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3):837–845.

Duval, R. A., Furceri, D., Hu, B., Jalles, J. T., and Nguyen, H. (2018). A Narrative Database of Major Labor and Product Market Reforms in Advanced Economies. IMF Working Papers 2018/019, International Monetary Fund.

Enikolopov, R., Makarin, A., and Petrova, M. (2020). Social media and protest participation: Evidence from russia. *Econometrica*, 88(4):1479–1514.

Freund, Y. and 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.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–74.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning.* MIT Press.

Hadzi-Vaskov, M., Pienknaguraand, S., and Ricci, L. A. (2021). The macroeconomic impact of social unrest. *IMF Working paper WP/21/135.*

Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning.* Springer Series in Statistics. Springer New York Inc., New York, NY, USA.

Hellwig, K.-P. (2020). Predicting fiscal crises: A machine learning approach. *mimeo.*

Hlatshwayo, S. and Redl, C. (2020). The macroeconomic relevance of social unrest. *forthcoming.*

Joseph, A. (2019). Parametric inference with universal function approximators. Bank of England working papers 784, Bank of England.

Kaminsky, G. L. and Reinhart, C. M. (1999). The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review*, 89(3):473–500.

Kingma, D. P. and Ba, J. (2017). Adam: A method for stochastic optimization.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc.

Madestam, A., Shoag, D., Veuger, S., and Yanagizawa-Drott, D. (2013). Do Political Protests Matter? Evidence from the Tea Party Movement*. *The Quarterly Journal of Economics*, 128(4):1633–1685.

Manacorda, M. and Tesei, A. (2020). Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567.

Moseley, M. W. (2015). Contentious engagement understanding protest participation in latin american democracies. *Journal of Politics in Latin America*, 7(3):3–48.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Racine, J. (2000). Consistent cross-validatory model-selection for dependent data: hv-block cross-validation. *Journal of Econometrics*, 99(1):39 – 61.

Redmon, J. and Farhadi, A. (2017). Yolo9000: Better, faster, stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525.

Shao, J. (1993). Linear model selection by cross-validation. *Journal of the American Statistical Association*, 88(422):486–494.

Strumbelj, E. and Kononenko, I. (2010). An efficient explanation of individual classifications using game theory. *Journal of Machine Learning Research*, 11.

Turchin, P. and Korotayev, A. (2020). The 2010 structural-demographic forecast for the 2010 to 2020 decade: A retrospective assessment. *PLoS ONE 15(8)*.

7 Appendix I: Machine learning models

A range of ML tools were compared for prediction performance. All models were compared using the expanding window approach described in the text, based on the AUC. We use the Python library of Pedregosa et al. (2011), scikit-learn.org, to implement these algorithms. This section draws heavily on the latter source, as well as Hastie et al. (2001).

7.1 Linear models: regularized logistic regression

The regularized logistic regression solves the following problem:

$$\min_{\beta,\alpha} \left\{ \frac{1-\rho}{2}||\beta||_2^2 + \rho||\beta||_1 + C\sum_{i=1}^{n} \left[\log\left(\exp\left(-y_i\left(X_i^T\beta+\alpha\right)\right)\right)+1\right] \right\}$$

This is a standard logistic regression where a penalty is applied to the size of the slope parameters, β. If $\rho = 1$ then this is the Lasso estimator for a logistic regression. If the $\rho = 0$, this is a Ridge logistic regression. Note that we assumed that $C = 1$. Given this formulation no hyper parameters require tuning and we simply ran the expanding window test for the two extreme values of ρ.

7.2 Neural Network

Neural networks take linear combinations of input variables, combine these using a nonlinear activation function (similar to a logistic regression), and then using the output as a derived input for prediction. That process can be repeated many times into layers of derived inputs. Here we outline the basic principles and direct the reader to Hastie et al. (2001) for an introduction or to Goodfellow et al. (2016) for a discussion of deep neural networks. We seek to minimize the loss function $L(y, \hat{y}, W) = -y\ln\hat{y} - (1-y)\ln(1-y) + \hat{\alpha}||W||_2^2$, where W is a matrix of weights or parameters, y is an observation of the target variable, $\hat{y}$ is a prediction for the target and α is a hyper parameter which was set to 0.0001. This loss function is minimized given the equations for $i = 1, ..., m$ layers of the network:

$$Z^{[i]} = W^{[i]}X + b^{[i]}$$

$$A^{[i]} = g(Z^{[i]})$$

$$Z^{[i+1]} = W^{[i+1]}X + b^{[i+1]}$$

$$A^{[i+1]} = g(Z^{[i+1]})$$

The weights matrix, W, is n_ixm, where n_i is the number of nodes used for layer i. Note the square brackets superscript indicates the position in the network rather than an exponent. The weights are solved using the gradient descent algorithm of Kingma and Ba (2017). A $m = 3$ layer network was assumed with grid search over the following choices for the number nodes in each layer:

- layer 1: **5**,10,20,30

- layer 2: **10**, 8, 5 ,1

- layer 3: **10**, 5, 1

This leads to 48 models whose forecasts are tested. The optimal parameters were: 5,10 and 10 for the three respective layers (highlighted in bold above).

7.3 Support vector machine

A support vector machine aims to use a linear model to separate points in the space of features that are associated with different classes (in our case unrest or no unrest). Specifically, the goal is to find a set of slope coefficients, w, and intercepts, b such that the prediction made by $sign\left(w^T\phi(x)+b\right)$ gives the correct classification for the target $y=\{1,-1\}$ where unrest is 1 and no unrest -1. We (mean-variance) normalise the set of predictors, prior to applying the algorithm so that all have comparable scale. Typically classification problems are not perfectly separable, some samples are allowed to be ζ_i from the boundary. The formal statement is:

$$\min_{w,b,\zeta} \frac{1}{2}w^T w + C\sum_{i=1}^{n}\zeta_i$$

$$s.t.\, y_i\left(w^T\phi(x)+b\right) \geq 1-\zeta_i$$

$$\zeta_i \geq 0, i=1,\ldots,n$$

Where $K(x,x')=\phi(x)^T\phi(x)=\exp\left(-\gamma||x-x'||^2\right)$. C controls the strength of the penalty on the boundary, with a low C leading a smoother decision boundary. γ affects how much a single observation can affect the shape of the boundary, a higher γ reduces the influence of a single observation v. a cluster of observations. We estimate the model on grid for $C=\{0.1,\mathbf{1},10,100,1000,10000\}$ and for $\gamma=\{\mathbf{0.001},0.01,0.1,1,10,100\}$- with the optimal parameters $\{C,\gamma\}=\{1,0.001\}$.

7.4 Tree based models

The tree based ensemble method of Random Forests was described in detail in the text. Here we note that we searched over a grid for the max depth of the trees cover $\{1,\mathbf{3},5\}$ and the maximum number of features from which to the bootstrapped sample Z covering $\{50,100,\mathbf{200}\}$ features. The optimal parameters were a max tree depth of 3 and max features of 200.

7.4.1 AdaBoost

AdaBoost is a competing ensemble method which performed well in our tests. AdaBoost takes a collection of weak learners (here short decision trees, or stumps) and re-weights them to produce a stronger ensemble learner. The re-weighting scheme puts higher weight on better performing models and on unexplained observations. The hyper parameters for this algorithm are the number of simple or short trees to consider and the learning rate which effectively amplified the role of the weights use to aggregate the simple models. Below is a description of how the

algorithm works:

Algorithm 2 AdaBoost

---

1. Initialize weigts for observations $w_i = 1/N, i = 1, ..., N$

2. For $m = 1$ to M:

    (a) Fit a decision tree, T_m to the training data, use the weights in step one to weight the prediciton associated with observation.

    (b) Compute

    $$err_m = \frac{\sum_{i=1}^{N} w_i I(y_i \neq T_m(x_i))}{\sum_{i=1}^{N} w_i}$$

    (c) Compute $\alpha_m = \ln\left((1 - err_m)/err_m\right)$

    (d) Set $w_i = w_i \exp\left(\alpha_m I(y_i \neq T_m(x_i))\right)$ for $i = 1, ..., N$

3. The update the output of the model using $G(x) = sign\left\{\sum_{m=1}^{M} \alpha_m T_m(x)\right\}$

4. Update the model using the learning rate: $T_t(x) = T_{t-1}(x) + \nu G(x)$

We considered a number of estimators ranging form 50 to 300 in increments of 25 and a learning rate of 0.1 to 0.6 in increments of 0.1. We found the optimal performance when the number of estimators is 50 and learning rate 0.1.

7.4.2 Gradient Boosted Trees

The final method we consider, which provided comparable performance to the Random Forest classifier, is Gradient Boosted Trees. Gradient boosting, similar to AdaBoost, builds strong learners from weak.

$$\hat{y}_i = F_M(x_i) = \sum_{i=1}^{M} \nu h_m(x_i)$$

Where M is the number of estimators, or simple/short decision trees. The recursive formulation is $F_m(x) = F_{m-1}(x) + \nu h_m(x)$, where we can think of $h_m(x)$ as a new tree added to the existing model and ν the learning rate, with the goal of minimizing the loss of the overall model:

$$\min_{h_m} \sum_{i=1}^{M} L(y_i, F_{m-1}(x) + \nu h_m(x))$$

The problem is initialized with F_0, which is a constant. In the case of a square loss function this would just be the mean of y. The loss function can be approximated by a first-order expansion:

$$L(y_i, F_{m-1}(x_i) + h_m(x_i)) \approx L(y_i, F_{m-1}(x_i)) + \nu h_m(x_i)\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F=F_{m-1}}$$

The quantity in square brackets is the derivative of L with respect to its second parameter, evaluated at $F_{m-1}(x_i)$, we denote by g_i. Note that the first term will be a constant. So we have:

$$h_m = \arg\min_h L_m \approx \arg\min_h \nu h_m(x_i)g_i$$

$h(x_i)$ will minimize h_m when it is chosen to be proportional to $-\nu g_i$. The above is applicable to the case of a

continuous target variable, for the case of classification the only alteration required is to map $F(x)$ into a probability, which can be done, for example, with a sigmoid function. The algorithm can be combined with bootstrap sampling of the features at each iteration, called sub sampling in the toolkit of Pedregosa et al. (2011).

We considered 4 hyper parameters. First, the number of estimators with a grid of $\{\mathbf{30}, 50, 100, 200\}$. Second, the learning rate with a grid of $\{\mathbf{0.1}, 0.2, 0.5\}$. Third, the subsample with a grid of $\{\mathbf{0.8}, 1\}$. Finally, a choice of $\mathbf{1}$ or 3 for the max depth of the trees. The optimal combination is $\{30, 0.1, 0.8, 1\}$, respectively.

8 Appendix II: Input Data and Aggregation scheme

	Variable Name, transformation	Aggregate Category	Source
0	AE Dummy	Dummy	Hellwig (2020)
1	EM Dummy	Dummy	Hellwig (2020)
2	LIC Dummy	Dummy	Hellwig (2020)
3	EM or Small State Dummy	Dummy	Hellwig (2020)
4	LIDC Dummy	Dummy	Hellwig (2020)
5	Small Developing State Dummy	Dummy	Hellwig (2020)
6	Non-SDS LIDC Dummy	Dummy	Hellwig (2020)
7	Non-SDS EM Dummy	Dummy	Hellwig (2020)
8	Low Income Dummy	Dummy	Hellwig (2020)
9	Lower-Middle Income Dummy	Dummy	Hellwig (2020)
10	Upper Middle Income Dummy	Dummy	Hellwig (2020)
11	High Income non-OECD Dummy	Dummy	Hellwig (2020)
12	High Income OECD Dummy	Dummy	Hellwig (2020)
13	Dummy: AFR Department	Dummy	Hellwig (2020)
14	Dummy: APD Department	Dummy	Hellwig (2020)
15	Dummy: EUR Department	Dummy	Hellwig (2020)
16	Dummy: MCD Department	Dummy	Hellwig (2020)
17	Dummy: WHD Department	Dummy	Hellwig (2020)
18	Dummy: South Asia	Dummy	Hellwig (2020)
19	Dummy: East Asia and Pacific	Dummy	Hellwig (2020)
20	Dummy: Asia and Pacific	Dummy	Hellwig (2020)
21	Dummy: LAC	Dummy	Hellwig (2020)
22	Dummy: North America	Dummy	Hellwig (2020)
23	Dummy: Americas	Dummy	Hellwig (2020)
24	Dummy: European Union	Dummy	Hellwig (2020)
25	Eur. and C. Asia excl. EU dummy	Dummy	Hellwig (2020)
26	Eur. and C. Asia	Dummy	Hellwig (2020)
27	MENA Dummy	Dummy	Hellwig (2020)
28	SSA Dummy	Dummy	Hellwig (2020)
29	Island Dummy	Dummy	Hellwig (2020)
30	Landlocked Dummy	Dummy	Hellwig (2020)
31	Dummy: Fuel exporter	Dummy	Hellwig (2020)
32	Exporter Classification	Dummy	Hellwig (2020)
33	Dummy: Small state	Dummy	Hellwig (2020)
34	Dummy: Micro state	Dummy	Hellwig (2020)
35	Dummy: Net creditor	Dummy	Hellwig (2020)
36	Dummy: Net debtor	Dummy	Hellwig (2020)
37	Dummy: Fragile state	Dummy	Hellwig (2020)
38	Dummy: Monetary Union	Dummy	Hellwig (2020)
39	Year Joining Mon. Union	Dummy	Hellwig (2020)
40	Dummy: HIPC country	Dummy	Hellwig (2020)
41	Dummy: Early HIPC	Dummy	Hellwig (2020)
42	Dummy: Late HIPC	Dummy	Hellwig (2020)
43	Dummy: HIPC pre-decision	Dummy	Hellwig (2020)
44	Dummy: VE-LIC	Dummy	Hellwig (2020)
45	Dummy: VE-LIC Com. Exporter	Dummy	Hellwig (2020)
46	Dummy: Fuel or Comm. exporter	Dummy	Hellwig (2020)
47	Dummy: Comm. exporter, EM/LIC	Dummy	Hellwig (2020)
48	Oil Price Inflation	Oil Inflation	Hellwig (2020)
49	Non-fuel Price pct Change	Inflation	Hellwig (2020)
50	Food Price Inflation	Food Inflation	Hellwig (2020)
51	Inflation	Inflation	Hellwig (2020)
52	Per Capita GDP Growth	GDP Growth	Hellwig (2020)
53	Exchange Rate Depreciation (pct)	Exchange Rate	Hellwig (2020)
54	Terms of Trade Inflation	Terms of Trade	Hellwig (2020)
55	Growth Deviation from 5y Avg	GDP Growth	Hellwig (2020)
56	Real Exchange Rate Growth	Exchange Rate	Hellwig (2020)
57	Nat. Disaster Impact Growth	Nautral Disasters	Hellwig (2020)
58	Nat. Disaster Impact Growth, lag	Nautral Disasters	Hellwig (2020)
59	Prim. Bal./GDP	Primary Balance	Hellwig (2020)
60	Prim. Bal./GDP, lag	Primary Balance	Hellwig (2020)
61	GG Exp./GDP	Fiscal Expeditures	Hellwig (2020)
62	GG Exp./GDP, lag	Fiscal Expeditures	Hellwig (2020)
63	GG Revenue /GDP	Fiscal Revenue	Hellwig (2020)
64	GG Revenue /GDP , lag	Fiscal Revenue	Hellwig (2020)
65	Net Aid/GDP	Aid	Hellwig (2020)
66	Net Aid/GDP, lag	Aid	Hellwig (2020)
67	GG Prim. Exp./GDP	Fiscal Expeditures	Hellwig (2020)
68	GG Prim. Exp./GDP, lag	Fiscal Expeditures	Hellwig (2020)
69	Remittances / GDP	Remittances	Hellwig (2020)
70	Remittances / GDP , lag	Remittances	Hellwig (2020)
71	Polity Score	Political	Hellwig (2020)
72	Polity Score, lag	Political	Hellwig (2020)
73	Exec. Term Remaining	Elections	Hellwig (2020)
74	Legislative Election Held Dummy	Elections	Hellwig (2020)
75	Exec. Election Held Dummy	Elections	Hellwig (2020)
76	Checks and balances index	Political	Hellwig (2020)
77	Oil Price Acceleration	Oil Inflation	Hellwig (2020)
78	Non-fuel Price Acceleration	Inflation	Hellwig (2020)
79	Food Price Acceleration	Food Inflation	Hellwig (2020)
80	CPI Acceleraion	Inflation	Hellwig (2020)
81	RGDP per Capita, first diff.	GDP Growth	Hellwig (2020)

82	Exchange Rate Acceleration	Exchange Rate	Hellwig (2020)
83	ToT Inflation Growth	Terms of Trade	Hellwig (2020)
84	Reserves Acceleration	Reserves	Hellwig (2020)
85	Real GDP Acceleration	GDP Growth	Hellwig (2020)
86	Real Exchange Rate Acceleration	Exchange Rate	Hellwig (2020)
87	Prim. Bal./GDP Change	Primary Balance	Hellwig (2020)
88	Prim. Bal./GDP Change, lag	Primary Balance	Hellwig (2020)
89	GG Exp./GDP Growth	Fiscal Expeditures	Hellwig (2020)
90	GG Exp./GDP Growth, lag	Fiscal Expeditures	Hellwig (2020)
91	GG Rev. /GDP Growth	Fiscal Revenue	Hellwig (2020)
92	GG Rev. /GDP Growth, lag	Fiscal Revenue	Hellwig (2020)
93	GG Interest Exp./GDP Growth	Interest Expense	Hellwig (2020)
94	GG Interest Exp./GDP Growth, lag	Interest Expense	Hellwig (2020)
95	Net Aid/GDP Growth	Aid	Hellwig (2020)
96	Net Aid/GDP Growth, lag	Aid	Hellwig (2020)
97	GG Prim. Exp./GDP, first diff.	Fiscal Expeditures	Hellwig (2020)
98	GG Prim. Exp./GDP, first diff., lag	Fiscal Expeditures	Hellwig (2020)
99	Remittances / GDP Growth	Remittances	Hellwig (2020)
100	Remittances / GDP Growth, lag	Remittances	Hellwig (2020)
101	Trading Partner Growth	Trading Partner Growth	Hellwig (2020)
102	Inflation, EOP	Inflation	Hellwig (2020)
103	Growth of EOP Inflation	Inflation	Hellwig (2020)
104	RGDP Growth Volatility	GDP Growth	Hellwig (2020)
105	ToT Volatility	Terms of Trade	Hellwig (2020)
106	Exchange Rate Volatility	Exchange Rate	Hellwig (2020)
107	Inflation Volatility	Inflation	Hellwig (2020)
108	Polity Score, 5y growth	Political	Hellwig (2020)
109	Polity Score Growth	Political	Hellwig (2020)
110	Polity Score Growth, lag	Political	Hellwig (2020)
111	Dummy: Monetary Union	Dummy	Hellwig (2020)
112	Bureaucracy Quality	Public Services	Hellwig (2020)
113	Corruption	Political	Hellwig (2020)
114	Net Aid/GDP 3y diff.	Aid	Hellwig (2020)
115	Prim. Bal./GDP 3y Change	Primary Balance	Hellwig (2020)
116	GG Prim. Exp./GDP, 3y Change	Fiscal Expeditures	Hellwig (2020)
117	Remittances / GDP 3y Growth	Remittances	Hellwig (2020)
118	GDP, USD, log	Size of Economy	Hellwig (2020)
119	Population 10y Growth	Population and urbanisation	Hellwig (2020)
120	Population 5y Growth	Population and urbanisation	Hellwig (2020)
121	Population	Population and urbanisation	Hellwig (2020)
122	Exchange Rate (EOP) Volatility	Exchange Rate	Hellwig (2020)
123	Oil Price Inflation, lag	Oil Inflation	Hellwig (2020)
124	Non-fuel Price pct Change, lag	Inflation	Hellwig (2020)
125	Food Price Inflation, lag	Food Inflation	Hellwig (2020)
126	Inflation, lag	Inflation	Hellwig (2020)
127	Per Capita GDP Growth, lag	GDP Growth	Hellwig (2020)
128	Exchange Rate Depreciation (pct), lag	Exchange Rate	Hellwig (2020)
129	Terms of Trade Inflation, lag	Terms of Trade	Hellwig (2020)
130	Real GDP , lag	GDP Growth	Hellwig (2020)
131	Inflation, EOP, lag	Inflation	Hellwig (2020)
132	Exchange Rate, end of period, lag	Exchange Rate	Hellwig (2020)
133	Oil Price 3y Inflation	Oil Inflation	Hellwig (2020)
134	Non-fuel Price 3y pct Change	Inflation	Hellwig (2020)
135	Food Price 3y Inflation	Food Inflation	Hellwig (2020)
136	Inflation, 3y Avg	Inflation	Hellwig (2020)
137	Per Capita GDP 3y Growth	GDP Growth	Hellwig (2020)
138	Exchange Rate 3y Depreciation	Exchange Rate	Hellwig (2020)
139	ToT Inflation 3y Growth	Terms of Trade	Hellwig (2020)
140	NGP Growth, 3y Geom. Avg	GDP Growth	Hellwig (2020)
141	Real PPP GDP/ US GDP	Size of Economy	Hellwig (2020)
142	Reported Social Unrest Index	Reported Social Unrest Index	Barrett et al (2020)
143	Reported Social Unrest Index, lag 1	Reported Social Unrest Index	Barrett et al (2020)
144	Reported Social Unrest Index, lag 2	Reported Social Unrest Index	Barrett et al (2020)
145	Reported Social Unrest Index, lag 3	Reported Social Unrest Index	Barrett et al (2020)
146	Reported Social Unrest Index, lag 4	Reported Social Unrest Index	Barrett et al (2020)
147	Reported Social Unrest Index, lag 5	Reported Social Unrest Index	Barrett et al (2020)
148	Country Size in Square Kilometers	Size of Economy	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
149	Internet Hosts Per Capita (Scaling: 0.000001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
150	Internet Users Per Capita Scaling: (0.000001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
151	Estimated Personal Computers Per Capita (Scaling: 0.0001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
152	% Annual Increase: Population Density (Scaling: 0.01)	Population and urbanisation	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
153	% Annual Increase: Imports Per Capita (Scaling: 0.01)	Imports	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
154	% Annual Increase: Exports Per Capita (Scaling: 0.01)	Exports	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
155	% Annual Increase: Televisions Per Capita (Scaling: 0.01)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
156	% Annual Increase: Daily Newspaper Circulation Per Capita (Scaling: 0.01)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
157	% Annual Increase: Primary School Enrollment Per Capita (Scaling: 0.01)	Education	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
158	% Annual Increase: Secondary School Enrollment Per Capita (Scaling: 0.01)	Education	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
159	% Annual Increase: University Enrollment Per Capita (Scaling: 0.01)	Education	Cross-National Time Series (CNTS) Data Archive (2019 Edition)

160	% Annual Increase: Percent Literate (Scaling: 0.01)	Education	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
161	% Annual Increase: Physicians Per Capita (Scaling: 0.01)	Access to Basic Services	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
162	Assassinations	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
163	General Strikes	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
164	Terrorism/Guerrilla Warfare	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
165	Government Crises	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
166	Purges	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
167	Riots	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
168	Revolutions	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
169	Anti-Government Demonstrations	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
170	Weighted Conflict Index	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
171	Registered Voters/Population (Scaling: 0.001)	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
172	Effectiveness of Legislature	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
173	Competitiveness of Nominating Process	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
174	Party Coalitions	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
175	Party Legitimacy	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
176	Size of Legislature/Number of Seats, Largest Party (Scaling: 0.01)	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
177	Televisions Per Capita (Scaling: 0.00001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
178	Daily Newspaper Circulation Per Capita (Scaling: 0.0001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
179	National Defense Expenditure Per Capita (Scaling: 0.01)	Fiscal Expeditures	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
180	Mobile Cellular Telephones Per Capita (Scaling: 0.00001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
181	All Telephones, including Cellular, Per Capita (Scaling: 0.00001)	Digital and Media	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
182	Physicians Per Capita (Scaling: 0.000001)	Access to Basic Services	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
183	Party Fractionalization Index (Scaling: 0.0001)	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
184	Type of Regime	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
185	Number of Coups d'Etat	Unrest Other	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
186	Number of Major Constitutional Changes	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
187	Head of State	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
188	Premier	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
189	Effective Executive (Type)	Political	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
190	Effective Executive (Selection)	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
191	Degree of Parliamentary Responsibility	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
192	Size of Cabinet	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
193	Number of Major Cabinet Changes	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
194	Changes in Effective Executive	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
195	Legislative Effectiveness	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
196	Legislative Selection	Legislative	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
197	Number of Legislative Elections	Elections	Cross-National Time Series (CNTS) Data Archive (2019 Edition)
198	Prison Population, Total	Prisons	World Prison Brief
199	Persons Held in Prisons, Rate Per 100,000 population	Prisons	United Nations Office on Drugs and Crime
200	Bureaucracy Quality	Public Services	International Country Risk Guide
201	Current Account Risk (% GDP)	GDP Growth	International Country Risk Guide
202	Current Account Risk (% Exports)	CAB_Risk_XGS	International Country Risk Guide
203	Civil Disorder	Unrest Other	International Country Risk Guide
204	Civil War	War	International Country Risk Guide
205	Composite Risk Rating	Risk Rating	International Country Risk Guide
206	Consumer Confidence	Confidence	International Country Risk Guide
207	Contract Viability	Public Services	International Country Risk Guide
208	Corruption	Public Services	International Country Risk Guide
209	Cross-Border Conflict	Public Services	International Country Risk Guide
210	Democratic Accountability	Political	International Country Risk Guide
211	Economic Risk Rating	Risk Rating	International Country Risk Guide
212	Ethnic Tensions	Ethnic and Religious Tensions	International Country Risk Guide
213	External Conflict	War	International Country Risk Guide
214	Financial Risk Rating	Risk Rating	International Country Risk Guide
215	Foreign Pressures	Political	International Country Risk Guide
216	Exchange Rate Risk	Exchange Rate	International Country Risk Guide
217	Government Cohesion	Government	International Country Risk Guide
218	Government Stability	Government	International Country Risk Guide
219	Internal Conflict	War	International Country Risk Guide
220	Investment Profile	Public Services	International Country Risk Guide
221	Law & Order	Public Services	International Country Risk Guide

222	Legislative Strength	Public Services	International Country Risk Guide
223	Military in Politics	Political	International Country Risk Guide
224	Payment Delays	Government	International Country Risk Guide
225	Political Risk Rating	Political	International Country Risk Guide
226	Popularsupport	Government	International Country Risk Guide
227	Poverty	Poverty	International Country Risk Guide
228	ReligiousTension	Ethnic and Religious Tensions	International Country Risk Guide
229	Repatriation	Population and urbanisation	International Country Risk Guide
230	Risk for Budget Balance	Primary Balance	International Country Risk Guide
231	Risk for Debt Service	Debt	International Country Risk Guide
232	Risk for Foreign Debt	Debt	International Country Risk Guide
233	Risk for GDP Growth	GDP Growth	International Country Risk Guide
234	Risk for Inflation	Inflation	International Country Risk Guide
235	Risk for International Liquidity	Exchange Rate	International Country Risk Guide
236	Risk for Per Capita GDP	GDP Growth	International Country Risk Guide
237	Socioeconomic Conditions	Socioeconomic Conditions	International Country Risk Guide
238	Terrorism	War	International Country Risk Guide
239	War	War	International Country Risk Guide
240	Prevalence of moderate or severe food insecurity	Poverty	World Bank Development Indicators
241	Multidimensional poverty headcount ratio	Poverty	World Bank Development Indicators
242	Multidimensional poverty index	Poverty	World Bank Development Indicators
243	Poverty gap at $5.50 a day	Poverty	World Bank Development Indicators
244	Multidimensional poverty intensity	Poverty	World Bank Development Indicators
245	Access to electricity	Access to Basic Services	World Bank Development Indicators
246	People using at least basic sanitation services	Access to Basic Services	World Bank Development Indicators
247	People with basic handwashing facilities including soap and water	Access to Basic Services	World Bank Development Indicators
248	People using safely managed sanitation services	Access to Basic Services	World Bank Development Indicators
249	Access to clean fuels and technologies for cooking	Access to Basic Services	World Bank Development Indicators
250	Community health workers	Access to Basic Services	World Bank Development Indicators
251	Current health expenditure per capita	Access to Basic Services	World Bank Development Indicators
252	Hospital beds	Access to Basic Services	World Bank Development Indicators
253	Life expectancy at birth, male	Access to Basic Services	World Bank Development Indicators
254	Life expectancy at birth, female	Access to Basic Services	World Bank Development Indicators
255	Nurses and midwives	Access to Basic Services	World Bank Development Indicators
256	People using at least basic drinking water services	Access to Basic Services	World Bank Development Indicators
257	Physicians	Access to Basic Services	World Bank Development Indicators
258	Gini index	Inequality	World Bank Development Indicators
259	Income share held by highest 10%	Inequality	World Bank Development Indicators
260	Income share held by lowest 10%	Inequality	World Bank Development Indicators
261	Life expectancy at birth, female less male	Inequality	World Bank Development Indicators
262	School enrollment, primary and secondary (gross), gender parity index (GPI)	Inequality	World Bank Development Indicators
263	Literacy rate, youth (ages 15-24), gender parity index (GPI)	Inequality	World Bank Development Indicators
264	Account ownership, poorest 40%	Inequality	World Bank Development Indicators
265	Unemployment, youth male less female	Inequality	World Bank Development Indicators
266	Population in the largest city	Population and urbanisation	World Bank Development Indicators
267	Urban population	Population and urbanisation	World Bank Development Indicators
268	Population in urban agglomerations of more than 1 million	Population and urbanisation	World Bank Development Indicators
269	International migrant stock	Population and urbanisation	World Bank Development Indicators
270	Unemployment rate (ILO)	Unemployment	World Bank Development Indicators
271	Unemployment rate, youth female	Unemployment	World Bank Development Indicators
272	Unemployment rate, youth male	Unemployment	World Bank Development Indicators
273	Drought	Nautral Disasters	EM-DAT
274	Earthquake	Nautral Disasters	EM-DAT
275	Epidemic	Nautral Disasters	EM-DAT
276	Extreme_Temp	Nautral Disasters	EM-DAT
277	Flood	Nautral Disasters	EM-DAT
278	Industrial_accident	Nautral Disasters	EM-DAT
279	Storm	Nautral Disasters	EM-DAT
280	neighbours reported social unrest	Neighbours reported social unrest	Authors' calculation from Barrett et al (2020)
281	World Uncertainty Index	Uncertainty	Ahir et al (2018)
282	World Uncertainty Index, Lag 1	Uncertainty	Ahir et al (2018)
283	World Uncertainty Index, Lag 2	Uncertainty	Ahir et al (2018)
284	World Uncertainty Index, Lag 3	Uncertainty	Ahir et al (2018)
285	World Uncertainty Index, Lag 4	Uncertainty	Ahir et al (2018)
286	World Uncertainty Index, Lag 5	Uncertainty	Ahir et al (2018)

287	Alcoholic Beverages, Tobacco, and Narcotics_MAX	Inflation	IMF CPI Database
288	Clothing and footwear_MAX	Inflation	IMF CPI Database
289	Communication_MAX	Inflation	IMF CPI Database
290	Consumer Price Index, All items_MAX	Inflation	IMF CPI Database
291	Education_MAX	Inflation	IMF CPI Database
292	Food and non-alcoholic beverages_MAX	Food Inflation	IMF CPI Database
293	Furnishings, household equipment and routine household maintenance_MAX	Inflation	IMF CPI Database
294	Health_MAX	Health Inflation	IMF CPI Database
295	Housing, Water, Electricity, Gas and Other Fuels_MAX	Utilities Inflation	IMF CPI Database
296	Miscellaneous goods and services_MAX	Inflation	IMF CPI Database
297	Recreation and culture_MAX	Inflation	IMF CPI Database
298	Restaurants and hotels_MAX	Inflation	IMF CPI Database
299	Transport_MAX	Transport inflation	Duval et al (2018)
300	Domestic finance - Overall	Structural Reforms	Duval et al (2018)
301	Domestic finance - S1 - Banking supervisions	Structural Reforms	Duval et al (2018)
302	Domestic finance - S1 - Credit controls	Structural Reforms	Duval et al (2018)
303	Domestic finance - S1 - Entry barriers	Structural Reforms	Duval et al (2018)
304	Domestic finance - S1 - Interest rate controls	Structural Reforms	Duval et al (2018)
305	Domestic finance - S1 - Privatization	Structural Reforms	Duval et al (2018)
306	Domestic finance - S1 - Security markets	Structural Reforms	Duval et al (2018)
307	External finance - Overall	Structural Reforms	Duval et al (2018)
308	External finance - S1 - Capital inflow restrictions	Structural Reforms	Duval et al (2018)
309	External finance - S1 - Capital outflow restrictions	Structural Reforms	Duval et al (2018)
310	External finance - S2 - Inflow restrictions - Bank credit	Structural Reforms	Duval et al (2018)
311	External finance - S2 - Inflow restrictions - Bond	Structural Reforms	Duval et al (2018)
312	External finance - S2 - Inflow restrictions - FDI	Structural Reforms	Duval et al (2018)
313	External finance - S2 - Inflow restrictions - Money market instrument	Structural Reforms	Duval et al (2018)
314	External finance - S2 - Inflow restrictions - Portfolio equity	Structural Reforms	Duval et al (2018)
315	External finance - S2 - Outflow restrictions - Bank credit	Structural Reforms	Duval et al (2018)
316	External finance - S2 - Outflow restrictions - Bond	Structural Reforms	Duval et al (2018)
317	External finance - S2 - Outflow restrictions - FDI	Structural Reforms	Duval et al (2018)
318	External finance - S2 - Outflow restrictions - Money market instrument	Structural Reforms	Duval et al (2018)
319	External finance - S2 - Outflow restrictions - Portfolio equity	Structural Reforms	Duval et al (2018)
320	Labor market - Overall	Structural Reforms	Duval et al (2018)
321	Labor market - S1 - Additional requirements for collective dismissal	Structural Reforms	Duval et al (2018)
322	Labor market - S1 - Firing costs	Structural Reforms	Duval et al (2018)
323	Labor market - S1 - Procedural inconvenience	Structural Reforms	Duval et al (2018)
324	Labor market - S1 - Redress measures	Structural Reforms	Duval et al (2018)
325	Labor market - S1 - Valid grounds	Structural Reforms	Duval et al (2018)
326	Product market - Overall	Structural Reforms	Duval et al (2018)
327	Product market - S1 - Electricity	Structural Reforms	Duval et al (2018)
328	Product market - S1 - Telecommunications	Structural Reforms	Duval et al (2018)
329	Product market - S2 - Electricity - Ownership	Structural Reforms	Duval et al (2018)
330	Product market - S2 - Electricity - Regulation	Structural Reforms	Duval et al (2018)
331	Product market - S2 - Electricity - Unbundling	Structural Reforms	Duval et al (2018)
332	Product market - S2 - Electricity - Wholesale	Structural Reforms	Duval et al (2018)
333	Product market - S2 - Telecommunications - Access	Structural Reforms	Duval et al (2018)
334	Product market - S2 - Telecommunications - Competition	Structural Reforms	Duval et al (2018)
335	Product market - S2 - Telecommunications - Ownership	Structural Reforms	Duval et al (2018)
336	Product market - S2 - Telecommunications - Regulation	Structural Reforms	Duval et al (2018)
337	Trade - Overall	Structural Reforms	Duval et al (2018)
338	Trade - S1 - Payment restrictions for international trade	Structural Reforms	Duval et al (2018)
339	Trade - S1 - Tariffs	Structural Reforms	Duval et al (2018)
340	Trade - S2 - Restrictions on payments for good imports	Structural Reforms	Duval et al (2018)
341	Trade - S2 - Restrictions on payments for service imports	Structural Reforms	Duval et al (2018)
342	Trade - S2 - Restrictions on receipts from good exports	Structural Reforms	Duval et al (2018)
343	Trade - S2 - Restrictions on receipts from service exports	Structural Reforms	Duval et al (2018)