

Article

Validating Game-Theoretic Models of Terrorism: Insights from Machine Learning

James T. Bang¹, Atin Basuchoudhary^{2,*}  and Aniruddha Mitra³¹ Department of Economics, St. Ambrose University, Davenport, IA 52803, USA; bangjames@sau.edu² Department of Economics and Business, Virginia Military Institute, Lexington, VA 24450, USA³ Economics Program, Bard College, Annandale-On-Hudson, NY 12504, USA; amitra@bard.edu

* Correspondence: basuchoudharya@vmi.edu

Abstract: There are many competing game-theoretic analyses of terrorism. Most of these models suggest nonlinear relationships between terror attacks and some variable of interest. However, to date, there have been very few attempts to empirically sift between competing models of terrorism or identify nonlinear patterns. We suggest that machine learning can be an effective way of undertaking both. This feature can help build more salient game-theoretic models to help us understand and prevent terrorism.

Keywords: machine learning; terrorism; game theory



Citation: Bang, J.T.; Basuchoudhary, A.; Mitra, A. Validating Game-Theoretic Models of Terrorism: Insights from Machine Learning. *Games* **2021**, *12*, 54. <https://doi.org/10.3390/g12030054>

Academic Editors: Joao Ricardo Faria, Daniel Arce and Ulrich Berger

Received: 28 April 2021

Accepted: 23 June 2021

Published: 30 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Game-theoretic models of terrorism are a useful tool in understanding the interactions between states and terrorist groups, the organization of terror groups, and the coordination of counterterrorism efforts [1]. These models provide insights and testable hypotheses. Yet, far too often, many of these hypotheses remain untested. Even when model-generated hypotheses are tested, the focus is on the effect of a particular theory-generated variable on, say, the likelihood of terrorism. This testing may explore causal channels. However, because empirical evidence using traditional econometric channels is not organized to check for relative salience, the importance of a correlation or even causal effect relative to other such effects is unknown. This inability of traditional econometric techniques to check for relative salience in an organized way makes it hard to sift among competing theoretical models. This sort of sifting is essential for policy. A nation plagued by terrorist attacks needs to know which theoretical model provides the largest counterterrorism impact.

Further, it is essential to know how a particular variable may affect terrorism. Game-theoretic models have a key strength. They show comparative static or even dynamic (particularly in evolutionary models) equilibrium shifts. Thus, variables that affect terrorism may do so in nonlinear ways. Traditional econometric tests focused on parametric point estimates are not built to pick up these equilibrium shifts. Nonlinearity, of course, can often be imposed in econometrics. However, this forces the researcher to guess where these nonlinearities may be (squaring the variable, for example, defines a particular shape on a relationship that may or may not be accurate).

Data problems also plague traditional econometric tests of game-theoretic models. Terrorism is, thankfully, rare. However, empirically, this requires heroic assumptions about the distribution of data when making inferences. Even without the rarity aspect, hypotheses testing for significance requires assumptions about the underlying distribution of data that are swept under the rug. Game theoretic models highlight strategic interaction between agents, which are often endogenous. Thus, assumptions about the distribution of data are necessary to estimate efficient and unbiased estimators. Then, there is the issue of model specification. Model specification is often subject to a researcher's explicit and implicit biases. All of this contributes to charges of p hacking in academic research [2].

We suggest that the full power of game theoretical insights can be validated by machine learning. This is particularly important for science in cases such as the study of terrorism, where randomized control trials are impossible or unethical. Therefore, a key contribution of this paper is to introduce the emerging methodology of machine learning to the game-theoretic study of terrorism that can, to a great extent, overcome the limitations of classical regression-based methods [3].

This paper will identify a methodology to identify a robust list of factors that contribute to an increased risk of terrorism and would inform the government on precisely what information to monitor in order to be able to anticipate terrorist events before they occur and would hence contribute to the design of counterterrorism policy at the strategic level. This set of variables can be the starting point for further causal analysis [4]. Our approach will rank variables by predictive importance. Counterterrorism policy is by definition something whose effect happens in the future. Thus, more predictively important variables can be better candidates for policy.

Predictively important variables are not necessarily causal. However, all causal variables should predict. We can identify variables that do not predict well. This reduces the likelihood that these variables are causal. Thus, theoretical models that suggest such variables matter for terrorism are less likely to be explanations for terrorism. To the best of our knowledge this sort of approach is new in the literature on terrorism.

Game-theoretic models predict such nonlinear relationships in comparative static settings. We use machine learning technology to develop partial dependence plots that let the data reveal how predictive variables affect terrorism. This feature makes machine learning an essential vehicle for exploring nonlinear relationships between a policy variable of interest and its effect on the likelihood of terrorism. Moreover, because our algorithms are theory-agnostic, we can let the data speak to actual relationships that can iteratively help us build better game-theoretic models.

We lay down some conceptual foundations about terrorism in Section 2. Section 3 describes the machine learning techniques we use. We describe our data in Section 4. We report our results in Section 5. In Section 6 we provide examples of how our results can be helpful for validating game theoretic models. Section 7 concludes.

2. Conceptual Foundations

The game-theoretic approach to terrorism tries to identify and deter terrorists through a cost-benefit lens that highlights the deep interaction between attacker and defender. Terrorism is a choice for successful rebellions (e.g., in Algeria, Israel, and Cyprus;) [5,6]. Deterrence involves greater policing/punishment and policies to increase the opportunity costs of terrorism at the tactical level [7,8]. However, the very act of deterrence elicits a response [9]. For example, attackers' and defenders' efforts may be complementary, which implies that improving military defense may be counterproductive [10].

The choice of terrorism is also a consequence of the nature of the target. Terrorists will substitute away from hard targets, suggesting that piecemeal policies that focus on some targets at the expense of others may be unproductive [11]. The nature of the target drives even the type of terrorist attack, such that harder targets elicit more suicide attacks in the context of a club goods model [12]. Moreover, increased military aid creates a moral hazard problem in recipient countries who now have an incentive to have terrorists attack them [13].

These lines of research show that terrorism is not a thing in itself; it is a tactical choice driven by context. Further, the relationship between attackers and defenders is constantly changing. A priori, there is no reason to believe that these changes have a linear pattern: Enders and Hoover, for example, empirically show a nonlinear relationship between income and terrorism [14]. However, despite the nonlinear relationships predicted by game-theoretic attacker-defender models, most empirical tests, if any, only provide information on the significance of point estimates. This is insufficient for policymakers since potential underlying nonlinearities may make average point estimates unhelpful. A

deer hunter shooting a foot to the left of the deer and a foot to the right of the deer but claiming he shot the deer on average is correct but will go hungry.

Current empirical research has tended to identify the “correlates” of terrorism and has largely failed to identify a consistent set of such correlations. Thus, predicting terrorist attacks has so far mainly been speculative. Machine learning algorithms can provide scientifically cross-validated predictions of the likelihood of a terrorist attack to provide national security agencies with an abbreviated, cross-validated list of variables (i.e., policy levers) that can best identify and hopefully deter terrorism. Machine learning techniques identify the most predictive variables among those. These algorithms then identify validated data-driven relationships between a predictively important covariate of terrorism and the likelihood of terrorism. This approach helps develop better models because they are theoretically agnostic. This agnosticism can help sift between theoretical models—a good model should be able to predict robustly. At the same time, predicting the likelihood of terrorist attacks provides meaningful intelligence for preventing terrorism.

3. Machine Learning

Machine learning (ML) methods are a growing set of methods for predicting and classifying various outcomes. These approaches have two applications: validating policy recommendations and testing theory [15]. Policymakers need to understand the potential effect of a policy before it is implemented, by definition, a matter of prediction. A theory, too, must be able to predict behavior. Machine learning is not a silver bullet, but it can help with these issues.

Further, the machine learning techniques we use do not require assumptions about the underlying distributions of the variables and the error terms. Thus, statistical issues arising out of problems such as endogeneity may be less relevant in these prediction models. For example, say we can identify a highly predictive variable, say X , for terrorism. The predictive value alone, shorn of endogeneity considerations, suggests that policy and academic research should focus on understanding the relationship between X and terrorism. This investigation would include how other variables may influence X as well. Thus, machine learning is a good place to start an investigation as well. Just because X predicts terrorism does not mean it is causal. However, if it is a good predictor then there must be something about X that deserves further scrutiny. By the same token, variables that do not predict terrorism can hardly be causal. A causal variable, by definition, should be predictive. Theoretical models that highlight variables that fail to predict are therefore unlikely to be good explanations for terrorism. This logic allows us to eliminate nonpredictive variables from consideration as casual factors. This process of reasoning provides a path for eliminating theoretical models that are unlikely to causally explain terrorism.

From a policy perspective, predictive analysis has a more direct affect. Say the predictive variable X is, upon further econometric analysis, is also found to be causal. Then, X can potentially be a policy lever because we know X predictably causes terrorism. Therefore, manipulating X can potentially reduce terrorism. Thus, machine learned prediction analysis can supplement econometric techniques for policy analysis.

Everything we noted above can be done using econometrics. However, econometrics requires assumptions about the underlying distribution of the variables. The concomitant endogeneity and specification problems and potential solutions are both susceptible to bias and a source of competing explanations for terrorism. For example, a particular theoretical model might suggest an empirical link between a variable and terrorism that can be tested. Such a test may even reveal a causal link with the right instrument. Yet, without a sense of the predictive salience of this link relative to other competing links we can have no idea whether this causal link is good explanation for terrorism. This is particularly an issue for game theoretic models because these by definition highlight endogenous strategic interactions. Machine learning models, by focusing on accurate prediction even in the presence of endogeneity are particularly suited for the empirical investigation of game theoretic models.

This paper suggests that validated ML techniques can help determine whether a particular theoretical model of terrorism has predictive salience relative to others. In the process, we address some problems inherent in interpreting machine-learned results.

We will build an empirical model using several parametric and nonparametric ML techniques (classical regression, Poisson regression, artificial neural network, regression tree, bootstrap aggregating, boosting, and random forest) to measure how and how well publicly available economic, geographic, and institutional variables *predict* the frequency and severity of terror attacks [16]. The first step in this process will be to identify the machine learning approach that best predicts terrorism.

Next, using the best technique, we will identify the most important variables for predicting terrorism. This process can help validate the predictive salience of a theoretical model relative to others.

Finally, we plot the partial dependency plots for terrorism to show how each variable impacts terrorism across the distribution of its values. This technique is important because game-theoretic analysis gives us reason to believe that many of the correlates of terrorism have nonlinear impacts. Partial dependence plots also help us interpret results more meaningfully.

ML techniques identify tipping points in the range of a particular variable that may place a country at a lower or higher risk of terrorism. We illustrate these tipping points using partial dependence plots, which show how the incidence and severity of terror attacks fluctuate across each variable's observed values. Further, by identifying the variables that have the *most* predictive power, we could help develop a framework to distinguish between competing theoretical explanations of terrorism. Suppose, for instance, political models of terrorism may suggest that terrorism may be a tactic employed by disenfranchised groups with little or no voice in government. In contrast, economic models may suggest that groups employ terrorism as a signal of credibility to gain a seat at the negotiating table against the regime when it divvies up rents from resource wealth. Suppose ML methodologies rank democracy as a better predictor of terrorism than primary commodities exports, for example. In that case, we can assume that the political model may be a better explanation of terrorism than the economic model, or vice versa. Moreover, this approach can eliminate correlates of conflict that do not predict terrorism. Presumably, correlates that do not predict well cannot be considered as variables that cause terrorism. Such culling also helps build better specified and more precise models.

Our ML approach will help us better understand causal patterns explaining terrorism. Moreover, we offer a better understanding of how to predict terrorism, which will help policymakers design counterterrorist policies. The remainder of this section outlines the prediction algorithms we use to predict the aggregate terror risk for a country. Readers who are familiar with these algorithms—or will be bored by a technical description of them!—may skip to the results section. Those looking for a more detailed description of the algorithms may consult their coverage by [16].

3.1. Classical and Other Regression Analysis

Using given data from a learning sample, $L = \{(y_1, x_1), \dots, (y_N, x_N)\}$, any prediction function, $d(x_i)$, maps the vector of input variables, x , into the output variable (the number of terror attacks), y . An effective prediction algorithm seeks to define parameters that minimize an error function such as the mean absolute deviations or mean squared error, over the predictions. In linear regression models, $d(x_i)$ is simply a linear function of the inputs. A linear model with the MSE error function yields the ordinary least squares (OLS) regression model:

$$R_{OLS}(d) = \frac{1}{n} \sum_{i=1}^N (y_i - d(x_i))^2,$$

where $d(x_i) = x_i\beta$ is a linear function of the inputs.

Although OLS can sometimes yield good predictions (on average, the *best* prediction among all linear models, in fact), it has some undesirable properties in the case of predicting

terror attacks. Specifically, since a large number of cases in our sample experience no terror attacks at all, while some of them experience very large numbers of attacks, we will expect the OLS model to predict *negative* numbers of terror attacks for some observations—which is nonsense.

As an alternative, one corrects this problem by estimating a Poisson regression, which will estimate the average number of terror attacks conditional on the inputs, x , to be an exponential function of a linear combination of the inputs expressed as:

$$\lambda = E(y | x) = e^{\beta x}.$$

This means that the probability of observing a specific number of terror attacks will be:

$$p(y|x) = \frac{e^{y\beta} e^{-e^{\beta x}}}{y!}.$$

The Poisson model then proceeds by estimating the parameters to maximize the likelihood function for this Poisson probability distribution.

While these more sophisticated regression methods successfully purge the bias from the individual parameter estimates that might result from overdispersion, they do so to the detriment of the model's overall predictive accuracy. Alternative approaches, which ensure a relatively high degree of accuracy while also avoiding nonsensical predictions, use nonparametric tree methods or combinations of trees to predict the number of terror attacks.

3.2. Artificial Neural Networks (ANNs)

A feedforward artificial neural network is a series of binary regression models connecting each of the K input variables to M hidden nodes, over which, in the case of a regression problem such as ours (as opposed to a classification problem in the case of a binary target variable), a linear regression connects the hidden nodes to the output we hope to predict in the final layer. The logistic function is the usual activation function in the first layer, but in general any sigmoid function will have the desired properties. In a regression problem such as ours, the final layer usually contains only one output; the same is true for classification problems involving a binary output. For classification problems involving multinomial outputs, there can be any number of outputs. Thus, this methodology is quite flexible. Hence, with one output node, an ANN estimates $K \cdot M$ parameters.

We present a diagram of a simple ANN for predicting terror attacks in Figure 1. In the figure, each connection corresponds to a weight for each *input* variable ($I1, \dots, I5$) and bias (constant) terms ($B1$ and $B2$) into the hidden nodes ($H1$ and $H2$), or into the output node ($O1$). In the diagram, we show a two-layer neural network (the inputs do not count as a layer) with five inputs, two hidden nodes, and a constant. The link function connecting the hidden layer to the outputs, and which is not explicitly shown, is linear.

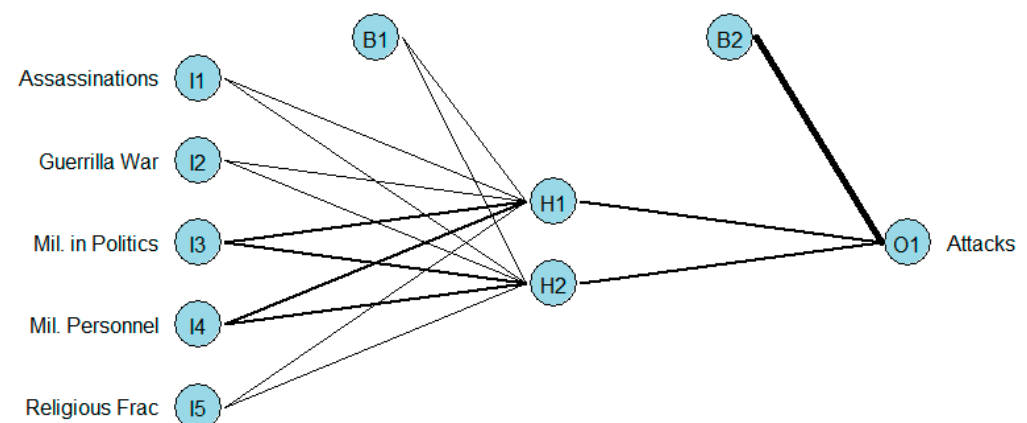


Figure 1. ANN Diagram Example.

Using a least-squares objective, the estimation of the ANN minimizes:

$$R(\alpha, \beta; \mathbf{x}) = \sum_{i=1}^N (y_i - f(\alpha, \beta; x_i))^2,$$

where $f(\alpha, \beta; x_i) = Z\beta$ connects the hidden layer to the output, and $Z = \frac{1}{1+\exp(x\alpha)}$ is the logit function connecting the inputs to the hidden layer. Using the first order conditions with respect to the parameters for the hidden layer, α , and the parameters for the output layer, β , the estimation finds the solution according to a gradient descent rule:

$$\beta_m^{r+1} = \beta_m^r - \sum_{i=1}^N \frac{\partial R}{\partial \beta_m^r} - \lambda \beta_m^r,$$

where λ is called the “weight decay” and acts as a penalty on the parameter and effectively restricts the parameters towards zero to avoid “overfitting” the model to the learning sample.

ANNs often perform well in situations where the interplay between input components is more important than any of their values. As such, they are often used in image and pattern recognition problems. We estimate the network using the *nnet* package implemented in R [17]. This implementation uses a single hidden layer (in which we used 100 nodes and 100 iterations). This work used all default options, save for specifying that the final layer should be linear. The initial weights were chosen randomly, and the goal function was the sum of the squared errors.

3.3. Regression Trees

Classification and regression trees (CART) diagnose and predict outcomes by finding binary splits in the input variables to optimally divide the sample into subsamples with successively higher levels of accuracy in the output variable, y . Therefore, unlike linear models, where the parameters are linear coefficients on each input variable, the parameters of the tree models are “if-then” statements that split the dataset according to the observed values of the inputs. We provide only a brief summary of tree construction as it pertains to our objectives [18].

More specifically, a tree, T , has four main parts:

1. Binary splits to splits in the inputs that divide the subsample at each node, t ;
2. Criteria for splitting each node into additional “child” nodes, or including it in the set of terminal nodes, T^* ;
3. A decision rule, $d(x)$, for assigning a predicted output value to each terminal node;
4. An estimate of the predictive quality of the decision rule, d .

The first step is achieved at each node by minimizing a measure of impurity. The most common measure of node impurity, and the one we use for our tree algorithms, is the mean square error, denoted $\hat{R}(d) = \frac{1}{n} \sum_{i=1}^N (y_i - d(x_i))^2$. Intuitively, this method searches for the cutoff in each input that minimizes errors, then selecting which input yields the greatest improvement in node impurity using its optimal splitting point.

Then, a node is declared to be terminal if one of the following conditions is met: (1) that the best split fails to improve the node impurity by more than a predetermined minimum improvement criterion; or (2) the split creates a “child” node that contains fewer observations than the minimum allowed (Note that there is a tradeoff here: setting lower values for the minimum acceptable margin of improvement or the minimum number of observations in a child node will lead to a more accurate prediction (at least within the sample the model uses to learn). However, improving the accuracy of the algorithm within the sample may lead to overfitting in the sense that the model will perform more poorly out-of-sample). At each terminal node, the decision rule assigns observations with a predicted outcome based on some measure of centrality. In the case of count (number of terror attacks or fatalities) or continuous (amount of property damage) outcomes, centrality is usually the mean of the observations conditional on reaching that node.

The predictive quality of the rule is also evaluated using the *mean square error*, $\hat{R}(d) = \frac{1}{n} \sum_{i=1}^N (y_i - d(x_i))^2$. This misclassification rate is often cross-validated by splitting the sample several times and re-estimating the misclassification rate each time to obtain an average misclassification of all of the cross-validated trees.

3.3.1. Boosting Algorithms

Iteratively re-estimating or combining ensembles of trees by averaging their predictions can often improve the accuracy of a tree algorithm. Boosting algorithms, bootstrap aggregating (bagging), and random forests all predict outcomes using ensembles of classification trees. The basic idea of these algorithms is to improve the predictive strength of a “weak learner” by iterating the tree algorithm many times by either modifying the distribution by reweighting the observations (boosting), randomly resampling a subset of the learning sample (bagging), or randomly sampling subsets of the input variables (random forest). These approaches then either classify the outcomes according to the outcome of the “strongest” learner once the algorithm achieves the desired error rate (boosting), or according to the outcome of a vote by the many trees (bagging).

Boosting has been proposed to augment the strength of a “weak learner” (an algorithm that predicts poorly) [19,20]. Specifically, for a given distribution \mathcal{D} of importance values assigned to each observation in L , and for a given desired error, \tilde{R} , and failure probability, φ , a *strong learner* is an algorithm that has a sufficiently high probability (at least $1 - \varphi$) of achieving an error rate no higher than \tilde{R} . A weak learner has a lower probability (less than $1 - \varphi$) of achieving the desired error rate. Boosting algorithms for classification create a set of M classifiers, $F = (f_1, \dots, f_M)$ that progressively reweight the importance of each observation based on whether the previous classifier predicted it correctly or incorrectly. Modifications of the boosting algorithm for classification have also been developed for regression trees [21,22].

Starting with a $\mathcal{D}_1 = (1/N, \dots, 1/N)$, suppose that our initial classifier, $f_1 = T$ (single-tree CART, for example), is a “weak learner” in that the misclassification rate, $\hat{R}(d)$ is greater than the desired maximum desired misclassification rate, \tilde{R} . Next, for all observations in the learning sample, recalculate the distribution weights for the observations as:

$$\mathcal{D}_2 = \frac{\mathcal{D}_1(i)}{Z_2} \times \begin{cases} \frac{\hat{R}_1(d)}{1 - \hat{R}_1(d)} & \text{if } d_1(x_i) = y_i \\ 1 & \text{otherwise} \end{cases},$$

where Z_m is a scaling constant that forces the weights to sum to one.

The final decision rule for the boosting algorithm is to categorize the outcomes according to $d(x) = \operatorname{argmax}_{y \in Y} \sum_{m: d_m(x)=y} \log\left(\frac{1 - \hat{R}_m(d)}{\hat{R}_m(d)}\right)$. Using this decision rule and its corresponding predictions, we calculate the estimate of the misclassification rate in the same way as in step (4) of the single tree algorithm.

3.3.2. Bootstrap Aggregating (Bagging)

The bagging method proposed by [23] takes random resamples, $\{L^{(M)}\}$, from the learning sample *with replacement* to create M samples using only the observations from the learning sample. Each of these samples will contain N observations—the same as the number of observations in the full training sample. However, in any one bootstrapped sample, some observations may appear twice (or more), others not at all. Note that the probability that a single observation is selected in each draw from the learning set is $1/N$. Hence, sampling with replacement, the probability that it is completely left out of any given bootstrap sample is $(1 - 1/N)^N$. For large samples this tends to $1/e$. The probability that an observation will be completely left out of all M bootstrap samples, then, is $(1 - 1/N)^{NM}$. The bagging method then adopts the rules for splitting and declaring nodes to be terminal described in the previous section to build M classification trees.

To complete steps (3) and (4), bagging needs a way of aggregating the information of the predictions from each of the trees. The way that bagging (and, as we will soon see, a random forest) does this for class variables is through *voting*. For *classification trees* (categorical output variables), the voting processes each observation through all of the M trees that was constructed from each of the bootstrapped samples to obtain that observation's predicted class for each tree. Note that the observations under consideration could be from the in-sample learning set or from outside the sample (the test set). The predicted class for the entire model, then, is equal to the mode prediction of all of the trees. For *regression trees* (continuous output variables), the voting process calculates the mean of the predicted values for all of the bootstrapped trees. Finally, the bagging calculates the redistribution estimate in the same way as it did for the single classification tree, using the predicted class based on the voting outcome.

3.3.3. Random Forests

Like bagging, a random forest is a tree-based algorithm that uses a voting rule to determine the predicted class of each observation. However, whereas the bagging randomizes the selection of the observations for each tree, a random forest may randomize over multiple dimensions of the classifier [24]. The most common dimensions for randomizing the trees are selecting the input variables for the node of each tree and the observations included for constructing each of the trees. We briefly describe the construction of the trees for the random forest ensemble below.

A random forest is a collection of tree decision rules, $\{d(x, \Theta_m), m = 1, \dots, M\}$, where Θ_m is a random vector specifying the observations and inputs that are included at each step of the construction of the decision rule for that tree. To construct a tree, the random forest algorithm takes to following steps:

- i Randomly select $n \leq N$ observations from the learning sample;
- ii At the "root" node of the tree, select $k \in K$ inputs from x ;
- iii Find the split in each variable selected in (ii) that minimizes the mean square error at that node and select the variable/split that achieves the minimal error;
- iv Repeat the random selection of inputs and optimal splits in (ii) and (iii) until some stopping criteria (minimum improvement, minimum number of observations, or maximum number of levels) is met.

The bagging method described in the previous subsection is in fact a special case of a random forest where, for each tree, Θ_m , of a random selection of $n = N$ observations from the learning sample with replacement (and each observation having a probability of being selected in each draw equal to $1/N$) and sets the number of inputs to select at each node, k , equal to the full length of the input vector, K so that all of the variables are considered at each node.

3.4. Validation and Testing of Predictive Accuracy

Once we have built our learning algorithm, the next issue is to evaluate the validity of our error estimates and the predictive strength of our models. Error estimates ($R[d]$) can sometimes be misleading if the model we are evaluating is overfitted to the learning sample. These error estimates can be tested out-of-sample or cross-validated using the learning sample.

To test the out-of-sample validity, we simply split the full dataset into two random subsets of *countries*: the first, known as the *learning sample* (or training sample) contains the countries and observations that will build the models; the second, known as the *test sample*, will test the out-of-sample predictive accuracy of the models. The out-of-sample error rates will indicate which models and specifications perform best, and will help reveal if any of the models are overfitted.

To validate the error rates, machine learning uses either hold-out validation or cross-validation. In our study, we have used hold-out validation, which involves training the models using one portion (in our case 70% selected at random) of the dataset. The algorithm

then tests the learned model by measuring the mean square error between the predicted value and the actual value in the 30% of the data unseen by it. A model with an acceptably low error rate in the sample unseen by it is presumably a good predictive model. This out of sample test also guards against overfitting. An overfitted model may be highly accurate in the learning sample but it would be unlikely to predict well in the test sample.

4. Data

As a first step in analyzing some preliminary data on terrorism, we have predicted the number of terror attacks using each of the seven models described above (OLS regression, Poisson regression, regression tree, random forest, bagging, and boosting). For our specification, we have included 69 input (or explanatory) variables that cover most of the ones discussed in Gassebner and Luechinger's survey of the empirical literature on conflict [25].

We measure our output (or "dependent") variable, Terror Attacks, as the total number of terror attacks in a country in the last five years. This variable comes from the Global Terror Database published by the University of Maryland and covers 1970–2014. When we combine all of the variables, our sample covers 1975–2014, since some entire data sources, such as the Database of Political Institutions, do not become available until 1975. To maintain the spirit of "prediction" in our model, we then consider our input ("explanatory") variables as five-year lagged averages of the preceding five years. Moreover, we only consider the variables at nonoverlapping five-year intervals so that none of the same information is contained in consecutive time intervals in our sample. In this sense, at any given point in time, policymakers will be able to use our model to predict whether a country will likely experience a greater or lesser number of terror incidents in the next five years. Moreover, this approach reduces the risk of endogeneity; the past can potentially affect the future, but it seems unlikely that the future can affect the past. In addition, this lagging reduces the risk of collider bias among the potential predictors if one were to interpret partial dependence plots causally. Collider bias happens when the target variable (Y, terrorism here) and a variable of theoretical interest (say T) affects a third variable, say X, in the model. In that case, if the researcher is interested in justifying a causal relationship between T and Y, X should be taken out of the model specification. Placing the target variable in the future helps justify that there can be no such relationship. We do *not* interpret our partial dependence plots causally.

From the Cross-National Time Series [26] we take the numbers of assassinations, demonstrations, government crises, guerrilla warfare incidents, purges, riots, and strikes as measures of underlying low-level social instability. We also take the number of cabinet changes and executive changes as measured of political instability, and the effectiveness of the legislature as a measure of political legitimacy.

From the Database of Political Institutions [27] we take the number of checks on power; executive and legislative indices of electoral competition; legislative, government, and opposition fractionalization indices; government Herfindahl index; and government polarization index as measures of the concentration (or not) of power and accountability (or not) within the government. We then include the changes in veto players, the existence of electoral fraud, executive tenure, the presence of a military executive, and political stability and executive power measures. Finally, we include plurality voting and proportional representation as indicators of structural differences in electoral rules.

Next, we take several indices of government quality from the International Country Risk Guide [28]. It is important to remember that, for each of the ICRG indices, a higher value always coincides with "better" outcomes on this dimension of institutional quality. For example, in the case of the "internal conflict" (or "external conflict") index, a higher value for the index somewhat counterintuitively corresponds to less conflict. The same can be said for "ethnic tensions", "religious tensions" and "military in politics"—in each of these cases, higher values relate to less of the (bad) thing that the variable name implies. That being said, we include the following indices from the ICRG: the bureaucratic quality and corruption indices as measures of the transparency of government; ethnic tensions,

external conflict, internal conflict, law and order, and religious tensions as measures of the levels of latent (or open) social hostility, and the government's ability to ease those hostilities; government stability and investment profile indices as measures of the government's credibility in carrying out stated policies and refraining from expropriation; and democratic accountability and military in politics indices as a measure of the legitimacy and responsiveness of the regime to the public's preferences. We also add the Polity2 index and regime durability from the Polity IV Project as additional measures of legitimacy and responsiveness.

As measures of economic and cultural divisions within society, we include measures of income inequality and ethnic and religious fractionalization. The former comes from the Standardized World Income Inequality Database [29]. The latter come from [30], which in turn come from the *Atlas Narodna Mira* [31].

Finally, we include numerous measures of economic human development from the World Development Indicators from the World Bank. They are: aid and development assistance; arms exports and imports; public education and health spending; female labor force participation; foreign direct investment (FDI); fuel exports; gross domestic product (GDP) per capita; government consumption; the stock of foreign born immigrants; infant mortality; the inflation rate in consumer prices; life expectancy; literacy; military expenditures; military personnel; population and its rate of growth; portfolio investment; primary, secondary, and tertiary school enrollment rates; social contributions; telephones per 100,000 people; the unemployment rate; urban population; and the youth dependency ratio.

Rather than exhaustively describing the distributional characteristics and justifying the inclusion of each variable, we kindly refer the reader to visit Gassebner and Luechinger's survey and the references therein to the various studies that have already provided such a description and justification [25]. For readers interested in some of the characteristics of the observed data in our sample, we have included the descriptive statistics for all 69 variables in Table 1.

We can see from the table that each of our explanatory variables has omitted values to varying degrees. The tree-based methods (single trees, boosting, bagging, and random forest) can automatically exploit the full information available by using surrogate information or using the median or mode at that branch of a tree as a best guess the value of a missing data point. Standard parametric methods (in our case Poisson regression and neural networks) do not do this automatically, and regression methods that do (such as full-information maximum likelihood), might do so in ways that give different imputations of the missing data.

To resolve this, we preprocess our data using random forest imputation. The basic idea is that we consider a covariate that does not have missing data (in our case conflict), and perform a random forest model to predict that variable (instead of the true variable of interest since that would be "cheating" for running the full model). Next, whenever the algorithm encounters a missing value at any tree node, the imputation substitutes the median or mode for that variable and continues with the subsequent splits. Therefore, the imputed values in each tree exploit the full complement of conditional distribution for that variable based on that tree. Averaging over all of the trees, we obtain imputed values for missing data points that uses as much relevant data about the conditional distribution of the variable as possible. It also has the advantage of creating imputed values that are naturally bounded by the domains of the observed data. Parametric methods such as multiple imputation estimate parameters based on an assumed distribution for the missing variables, and depending on the sensitivity of the parameters and the distributions of the covariates, may lead to extreme values outside of the logical bounds for a given variable (e.g., negative income).

Table 1. Variables and descriptive statistics.

Variable	Source	Obs	Mean	Std. Dev.	Min.	Max.
Terror Attacks	GTD	6411	86.88	375.81	0	10,701
Assassinations	CNTS	5318	0.21	0.84	0.00	18.50
Cabinet Changes	CNTS	5310	0.44	0.37	0.00	3.50
Demonstrations	CNTS	5318	0.52	1.15	0.00	14.00
Effectiveness of Leg.	CNTS	5297	1.74	0.94	0.00	3.00
Executive Changes	CNTS	5310	0.19	0.28	0.00	3.00
Government Crises	CNTS	5318	0.13	0.27	0.00	2.67
Guerrilla Warfare	CNTS	5318	0.12	0.32	0.00	2.60
Purges	CNTS	5318	0.03	0.13	0.00	2.50
Riots	CNTS	5318	0.31	1.05	0.00	18.20
Strikes	CNTS	5318	0.12	0.34	0.00	3.40
Changes in Veto Players	DPI	4838	0.12	0.15	0.00	1.00
Checks on Power	DPI	4831	2.52	1.60	1.00	17.00
Exec. Electoral Comp.	DPI	4850	5.15	2.08	1.00	7.00
Executive Years in Office	DPI	4859	7.93	7.68	1.00	45.00
Electoral Fraud	DPI	4214	0.14	0.32	0.00	1.00
Government Frac	DPI	4428	0.19	0.25	0.00	1.00
Government Herfindahl	DPI	4428	0.82	0.25	0.02	1.00
Government Polarization	DPI	4673	0.36	0.69	0.00	2.00
Legislative Frac.	DPI	4419	0.46	0.30	0.00	1.00
Leg. Electoral Comp.	DPI	4855	5.41	2.00	1.00	7.00
Military Executive	DPI	4856	0.21	0.39	0.00	1.00
Opposition Frac	DPI	3362	0.45	0.27	0.00	1.00
Plurality Voting	DPI	3877	0.68	0.46	0.00	1.00
Proportional Rep.	DPI	3474	0.58	0.49	0.00	1.00
Bureaucratic Quality	ICRG	3376	2.11	1.19	0.00	4.00
Corruption	ICRG	3376	3.08	1.35	0.00	6.00
Democratic Accountability	ICRG	3376	3.64	1.62	0.00	6.00
Ethnic Tensions	ICRG	3376	3.91	1.44	0.00	6.00
External Conflict	ICRG	3376	9.48	2.22	0.00	12.00
Government Stability	ICRG	3376	7.45	2.10	1.00	11.50
Internal Conflict	ICRG	3376	8.61	2.62	0.03	12.00
Investment Profile	ICRG	3376	6.94	2.34	0.08	12.00
Law and Order	ICRG	3376	3.60	1.48	0.25	6.00
Military in Politics	ICRG	3376	3.66	1.80	0.00	6.00
Religious Tensions	ICRG	3376	4.54	1.35	0.00	6.00
Polity2	Polity IV	4520	1.16	7.26	−10.00	10.00
Regime Durability	Polity IV	4569	23.99	28.73	0.00	198.00
Ethnic Fractionalization	Reynal-Querol	4749	0.45	0.28	0.01	0.96
Religious Fractionalization	Reynal-Querol	4749	0.28	0.23	0.00	0.78
Income Inequality (Gini)	SWIID	3350	38.52	9.87	16.49	69.35
Area	WDI	6110	682,865	1,717,163	2	16,400,000
Off. Aid & Dev. Assistance	WDI	4045	0.08	0.11	−0.01	0.76
Arms Exports	WDI	1703	0.01	0.08	0.00	1.50
Arms Imports	WDI	3976	0.04	0.12	0.00	3.32
Education Spending	WDI	3436	4.45	2.32	0.59	44.30
Foreign Direct Investment	WDI	4602	2.80	4.72	−32.30	72.50
Female Labor Force Part.	WDI	3293	50.12	17.55	9.20	90.80
Fuel Exports	WDI	3875	16.82	28.33	0.00	100.00
GDP per Capita	WDI	4807	9560.35	16,016.19	65.64	141,000.00
Government Consumption	WDI	4538	16.47	6.87	3.37	84.50
Health Spending	WDI	2647	3.48	2.21	0.01	18.36
Immigrant Stock	WDI	4975	8.07	13.75	0.03	86.80
Infant Mortality	WDI	5103	48.11	40.72	2.18	174.00
Inflation	WDI	4168	32.94	254.00	−17.60	6522.40
Life Expectancy	WDI	5074	64.79	10.58	24.30	82.50
Literacy Rate	WDI	1549	73.42	23.01	10.90	100.00
Military Expenditures	WDI	2995	2.74	3.03	0.09	48.60
Military Personnel	WDI	3092	1.88	2.23	0.06	35.80
Population	WDI	5190	30.94	116.87	8.82	1316.00
Population Growth	WDI	5190	1.80	1.44	−4.84	15.50
Portfolio Investment	WDI	4000	0.01	0.16	−0.02	4.88
Primary Enrollment	WDI	4763	97.05	22.35	15.80	208.00
Secondary Enrollment	WDI	4407	60.84	33.35	2.13	155.60
Social Contributions	WDI	1203	17.11	15.02	0.00	59.97
Telephones	WDI	5127	14.70	18.58	0.01	103.42
Tertiary Enrollment	WDI	4135	18.62	19.36	0.00	99.20
Unemployment	WDI	3007	9.03	6.78	0.20	59.50
Urban Population	WDI	5190	50.33	24.51	4.18	100.00
Youth Dependency	WDI	5000	62.07	23.94	19.44	114.40

5. Results

5.1. Predictive Quality

Table 2 reports the predictive quality of each of the models using the 70 variables. The best models we see to predict the overall number of terror attacks are the single regression tree, random forest, and bagging predictors, which reduce the overall MSE in the learning sample by about 64%, and 63%, and 59%, respectively, compared to the unconditional

sample mean. An average of all of the models' predicted values (which sometimes provides a better prediction, especially in cases of classification) improves the MSE by about 49%. In comparison, OLS regression improves the MSE by about 26%. However, as we might expect, the trees that use random bootstrapping (bagging and random forest) predict considerably better out of sample, with a test sample MSE reduction of 71% and 70% of the total MSE, respectively. Of particular interest here is the fact that these models achieve a significant reduction in the MSE despite the exclusion of the lagged number of terror attacks in our model since the pre-existing level of violence has been shown to be one of the strongest predictors of current and future violence in studies of conflict [32]. We exclude lagged terror attacks because we are partly looking to predict (a reason to include), but also looking to select a model to build theories and test causal effects (subsequent analyses). Lagged attacks would improve the prediction but would explain *so much* of the variation that we are not left with much to select a model on

Table 2. MSEs for the various learning models.

	Learning Sample		Test Sample	
	MSE	% Decrease	MSE	% Decrease
OLS Regression	107,708.05	25.71%	98,119.17	26.12%
Poisson Regression	151,539.85	−4.52%	139,385.78	−4.96%
Neural Network	144,695.12	0.20%	132,389.28	0.31%
Regression Tree	52,038.41	64.11%	80,182.62	39.62%
Boosting Predictor	141,677.19	2.28%	129,790.58	2.27%
Bagging Predictor	59,866.71	58.71%	40,202.12	69.73%
Random Forest	54,271.19	62.57%	38,504.85	71.01%
Average of All Predictors	74,564.39	48.57%	76,391.30	42.48%
Total MSE	144,987.24		132,802.82	

It is worth noting that the Poisson regression model, which tends to yield more valid estimates of causal effects, actually *increases* the MSE of the predictor compared to a prediction based on the simple sample mean. This is not quite the case for the neural network model, but we can see that the neural network and boosting models predict relatively poorly both in and out of sample.

5.2. Variable Importance

Table 3 reports the variable importance levels (measured as the percentage of the total reduction in MSE that is attributed to that variable) based on the single regression tree, boosting, bagging, and random forest models, which predicts conflict the best, although different algorithms or different runs of the same algorithm may identify different sets of predictors [15]. Theoretically agnostic algorithms may choose a predictive variable one time and another at a different time if they are predictive substitutes. The risk for this happening is reduced for algorithms such as random forests, bagging, or boosting because the algorithm learns by taking multiple subsamples and averaging the results. We take this one step further by averaging the variable importance results across several algorithms to give us a sense of confidence in the stability of the variable importance ranking.

Table 3. Variable importance rankings.

Variable	Tree	Bagging	Boosting	Forest	Average
Assassinations	7.618	24.930	62.966	12.388	14.979
Guerrilla War	2.677	10.735	30.698	9.436	7.616
Military Personnel	15.482	4.166	0.000	2.555	7.401
Religious Frac	12.386	4.761	0.000	3.218	6.788
Military Politics	12.682	1.913	1.082	3.529	6.042
Health Spending	3.765	4.390	2.499	3.548	3.901
Year	1.882	5.704	0.000	3.888	3.825
Population	0.947	3.568	0.394	5.562	3.359
Exec Yrs in Office	6.441	1.940	0.000	1.315	3.232
Fuel Exports	6.193	1.455	0.000	1.227	2.958
Dem Accountability	5.222	1.243	0.000	1.411	2.625
Effectiveness of Leg	0.000	3.104	0.000	3.041	2.048
Aid & Assistance	2.528	0.973	0.000	1.826	1.775
Gini	0.981	2.106	0.000	2.083	1.723
Tertiary Enrollment	2.053	0.752	0.000	2.023	1.609
Female LFPR	0.000	1.226	2.361	3.213	1.480
Portfolio Investment	0.000	2.695	0.000	1.600	1.432
Area	1.858	1.194	0.000	0.837	1.297
Arms Imports	1.425	1.402	0.000	1.009	1.279
Strikes	0.662	1.369	0.000	1.711	1.247
Ethnic Tension	0.733	0.467	0.000	2.409	1.203
Checks	1.702	1.248	0.000	0.615	1.188
Internal Conflict	0.969	0.125	0.000	2.257	1.117
Telephones	1.882	0.435	0.000	0.697	1.005
Law Order	1.322	0.560	0.000	0.966	0.950
GDP pc	0.235	0.842	0.000	1.736	0.938
Urban Population	1.710	0.404	0.000	0.645	0.920
Ethnic Frac	0.469	1.336	0.000	0.885	0.897
Polity 2	0.000	1.355	0.000	1.244	0.866
Investment Prof	0.321	1.019	0.000	1.169	0.836
Legislative Frac	1.425	0.532	0.000	0.549	0.835
Riots	0.307	0.729	0.000	1.382	0.806
Primary Enrollment	1.425	0.170	0.000	0.687	0.761
Arms Exports	0.000	1.233	0.000	0.950	0.728
Demonstrations	0.179	0.464	0.000	1.511	0.718
Unemployment	0.000	0.813	0.000	1.310	0.708
Religious Tension	0.000	0.300	0.000	1.601	0.634
Infant Mortality	0.000	0.773	0.000	1.046	0.606
Secondary Enrollment	0.000	0.848	0.000	0.765	0.537
Immigrant Stock	0.016	0.672	0.000	0.816	0.501
Reg Durability	0.248	0.547	0.000	0.609	0.468
Gov Consumption	0.000	0.369	0.000	0.824	0.398
Gov Stability	0.618	0.089	0.000	0.441	0.383
Corruption	0.075	0.472	0.000	0.584	0.377
Life Expectancy	0.000	0.242	0.000	0.889	0.377
Youth Dependency	0.000	0.286	0.000	0.842	0.376
FDI	0.000	0.372	0.000	0.719	0.363
Fraud	0.346	0.106	0.000	0.421	0.291
Opposition Frac	0.000	0.304	0.000	0.560	0.288
Inflation	0.207	0.326	0.000	0.300	0.278
External Conflict	0.259	0.140	0.000	0.428	0.276
Bureaucratic Qual	0.000	0.277	0.000	0.501	0.259
Leg. Elec. Comp.	0.248	0.192	0.000	0.295	0.245
Exec. Elec. Comp.	0.000	0.262	0.000	0.465	0.242
Literacy Rate	0.000	0.143	0.000	0.574	0.239
Population Growth	0.000	0.394	0.000	0.302	0.232
Proportional Rep	0.000	0.368	0.000	0.326	0.231
Social Contributions	0.000	0.173	0.000	0.515	0.229
Military Expend	0.167	0.212	0.000	0.277	0.219
Purges	0.331	0.051	0.000	0.179	0.187
Education Spending	0.000	0.243	0.000	0.304	0.182
Military Exec	0.000	0.085	0.000	0.316	0.134
Government Herfindahl	0.000	0.082	0.000	0.235	0.106
Plurality Voting	0.000	0.093	0.000	0.196	0.096
Gov Polarization	0.000	0.063	0.000	0.198	0.087
Government Frac	0.000	0.093	0.000	0.087	0.060
Cabinet Changes	0.000	0.028	0.000	0.085	0.038
Changes in Vetoes	0.000	0.065	0.000	0.018	0.028
Executive Changes	0.000	0.020	0.000	0.044	0.021
Government Crises	0.000	-0.050	0.000	-0.191	-0.080

Here, we see that the first five variables in the list account for close to one-third (about 31 percent) of the overall improvement in the random forest model's MSE. We also see that the single strongest predictor of current levels of terrorism is a history of assassinations in that country, which accounts for about 12% of the total reduction in the MSE in the random forest model, and 25% of the reduction for the bagging model and 63% of the reduction for the boosting model. The second strongest predictor, guerrilla war, accounts for about 10%

of the MSE reduction for the bagging and forest models and over 30% of the decrease for the boosting model.

After regime-directed violence, two of the following three strongest predictors involve the extent to which the military engages with everyday life and politics. Military personnel and the military in politics index account for almost 15% of the reduction in MSE combined, on average (slightly more in the single tree, somewhat less in the bagging and forest models, and not in the boosting algorithm). In between these measures of military engagement, we see religious fractionalization to account for about 7% of the variation on average. Rounding out the top ten predictors are health spending (3.9% of the MSE), time trend (3.8%), population (3.6%), executive tenure (3.2%), and fuel exports (2.6%).

At this point, our algorithmic approach suggests we have a group of variables that predict terrorism quite well. Moreover, we have identified the top predictors of terrorism. The reader will note that many of the variables identified by the literature do not have predictive salience [25]. Indeed, many of the variables highlighted in the literature, such as investment profile, bureaucratic quality, or religious tensions, have very little predictive salience. This culling helps us identify the kinds of theoretical models that can help us better understand terrorism. For example, the joint importance of guerilla war and military personnel is quite high and suggests that terrorism may be best understood as a tactical choice in asymmetric warfare rather than an outcome of institutional deficiencies in bureaucratic quality or lack of economic opportunity. This sort of explanation lends credence to the argument that a war on terror is strategically empty—just as a war on the blitzkrieg or the pincer movement, both tactical choices, would be strategically empty. However, such explanations are also predicated on the nature of the relationships between the top predictors and terrorist attacks. We turn to identify just such relations next, highlighting a methodological approach that is particularly in tune with the nonlinear relationships predicted by game-theoretic models.

5.3. The Nonlinear Relationship between Greater Security and Terrorism

The next step is to analyze *how* each of the variables impacts aggregate terror risk. To do this, we use a *partial dependence plot* mapping the possible values of the input variable of interest onto the observed incidence of terror attacks. Partial dependence plots display the marginal effect of variable x_k conditional on the observed values of all of the other variables, $(x_{1,-k}, x_{2,-k}, \dots, x_{n,-k})$. Specifically, it plots the graph of the function:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n f(x_k, x_{i,-k}),$$

where the summand, $f(x_k, x_{i,-k})$, is simply the observed outcome of the number of terror attacks.

This section focuses on three partial dependence plots that highlight game-theoretic models of terrorism that suggest that any fundamental understanding of terrorism should be understood as a tactical choice by rebel organizations.

Figure 2 shows that guerrilla warfare increases terrorism. While there may be some overlap between guerrilla warfare and terrorism, agencies that make national security policies tend to define them as distinct phenomena. Hence, in some cases, we might think of terrorism and guerrilla warfare as different tactics employed by rebel groups towards similar ends [33]. Moreover, guerilla warfare predicts terrorism five years out.

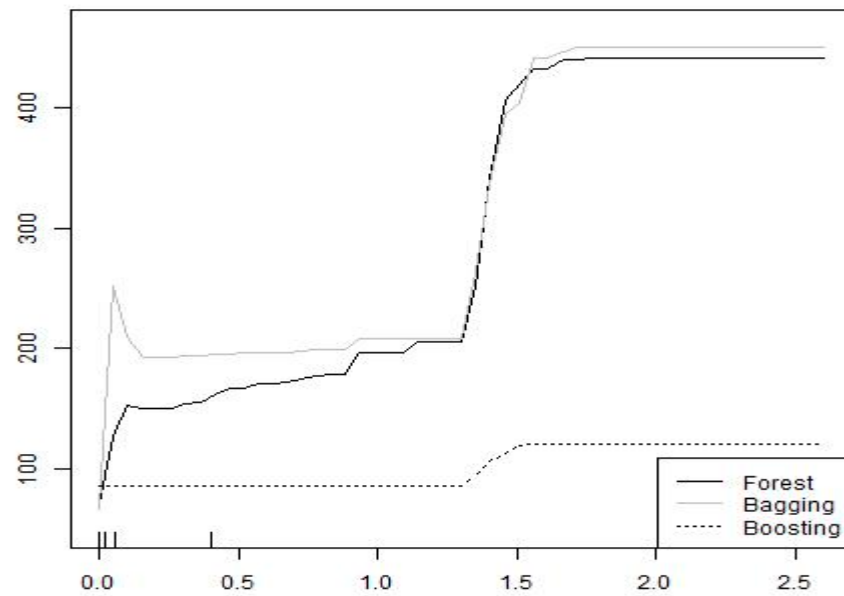


Figure 2. Partial dependence plot: guerrilla warfare.

Thus, in Figure 3, more military personnel also translate into more terror attacks on average, though these averages mask a u-shaped relationship. Last, in Figure 4, we notice that increased military involvement *in politics* reduces aggregate terror risk. Taken together, and in the absence of the predictive salience of such institutional variables such as bureaucratic quality and investment profile that capture elements of state capacity, we can grope toward a model of terrorism rooted in the understanding of a specific kind of state capacity. The nonlinear relationships embedded in this understanding suggest that game-theoretic models where equilibrium switching is possible due to interactions between agents are better suited than traditional neoclassical utility maximization approaches.

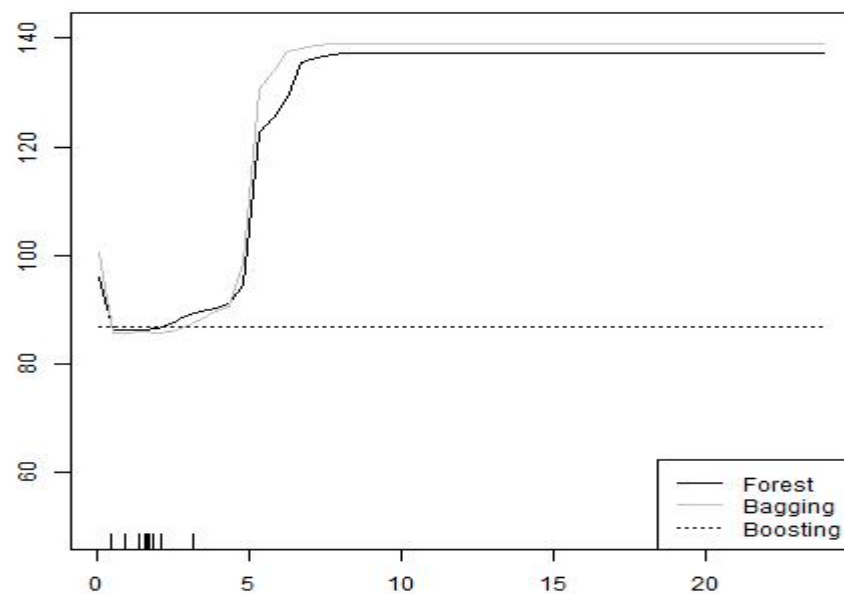


Figure 3. Partial dependence plot: military personnel.

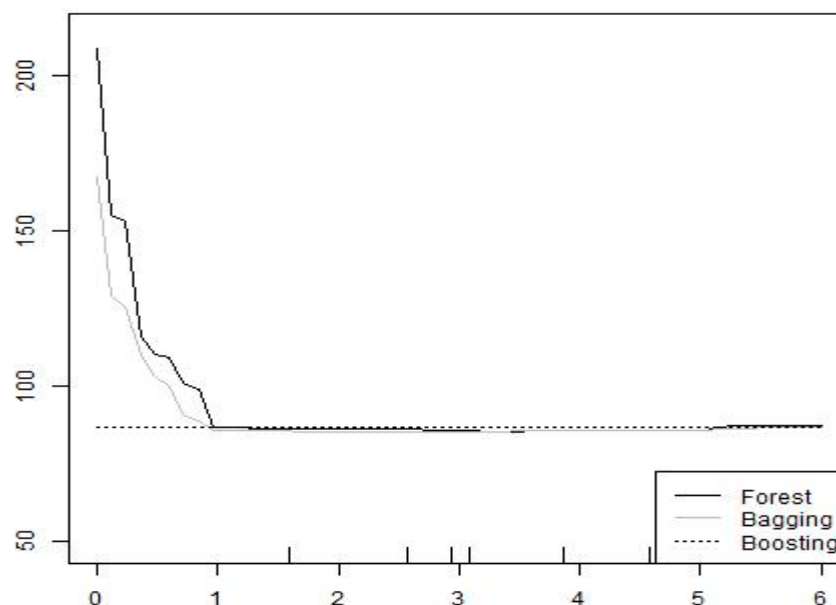


Figure 4. Partial dependence plot: military in politics.

State capacity (or the lack thereof) is a reasonably standard explanation for conflict [34]. The theoretical basis for an empirical understanding of this relationship lies in three concepts: military capacity, administrative quality, and institutional coherence.

Ref. [35] suggests that military personnel and expenditures, bureaucratic quality measures, and popular institutional measures such as polity or those reported by ICRG have construct and theoretical validity as a measure of the three elements of state capacity. We have all these variables as part of our predictive algorithm. Nevertheless, of these three, it appears that military capacity is most salient for understanding terrorism. Thus, our algorithm has been pretty specific about what kind of theoretical models are more likely explanations for terrorist attacks. This suggests that models better understand why terrorism happens [13]. The how matters as well.

There is a clear equilibrium switch in the number of terrorist attacks as guerilla warfare intensifies. However, there is an optimum level of intensity beyond which the number of terrorist attacks is stable. Again, this is the result suggested by game-theoretic models where equilibrium switches can be, for example, a consequence of changes in the payoffs. An econometric point parametric estimate would never capture these breakpoints unless, out of sheer coincidence, the researcher imposes the assumption of such a breakpoint. However, point estimates can be particularly misleading, for example, in the case of military personnel. An average effect captured in a point estimate would merely show a positive relationship, rather than the nuance where (initially at least) increasing military personnel reduces terrorism, thus suggesting a cost-minimizing optimum amount of military personnel. Nevertheless, we also have the somewhat counterintuitive but ultimately plausible result that hardening targets by increasing the number of military personnel elicits more terrorist attacks than substitute attacks away from these targets. This sort of result is reminiscent of security dilemmas rather than Beckerian policing models.

On the other hand, we cannot completely throw out institutional coherence as a predictor of terrorism. Military dictatorships can control terrorist attacks better. This result provides an interesting counterpoint to the argument that military regimes are more vulnerable to terrorism [36].

6. Game-Theoretic Model Validation

Others have suggested that machine learning can help validate theoretical models because they are designed to test whether a model is predictive or not [15,37]. A good theoretical model should be able to predict behavior.

Standard econometric approaches to testing models are particularly fraught when it comes to testing game-theoretic models because endogeneity is a feature rather than a bug in game-theoretic models. For example, terrorists respond to counterterrorism by changing their behavior, which in turn suggests changes in counterterrorism. Thus, any econometric approach to terrorism must be cautious to avoid endogeneity-driven estimation biases. Many of these methods reduce the predictive value of a model (for example, many causal studies have very low R-squares). Yet, as we noted above, a good theoretical model should also be able to predict. Predictive machine learning can help determine whether a causal variable is also predictive. A causal variable that is also predictive can help convince academics and policymakers of the salience of a theoretical model.

Partial dependence plots can capture equilibrium shifts to capture comparative static effects of game-theoretic models. We discuss this aspect quite extensively in the previous section. However, variable importance can help us sift through models of terrorism to identify more predictive variable specifications. We highlight three examples within subsets of the game-theoretic literature to emphasize this point.

One strand of the game theoretic literature focuses on group cohesion. Future uncertainty generated by increased counter terrorism can lead to rebel group splintering, thereby increasing the risk of terrorism as these splinter groups jockey for survival [38]. Figure 3 highlights just such a result; an increase in military personnel does indeed predict an increase in the number of terrorist attacks. Further, the first two most predictive variables, guerrilla warfare and assassinations, also predict an increase in the number of terrorist attacks. Guerrilla warfare and assassinations also point to significant political uncertainty. This suggests that political uncertainty may be an important predictor of terrorist attacks, possibly by affecting group cohesion. These findings would suggest a deeper, and causal, dive into understanding how rebel group splintering in the face of political uncertainty may affect terrorism. That is to say, machine-learning can be a first step toward finding explanations of terrorism in conjunction with game-theoretic models and causal econometric analysis.

Counterterrorism efforts require global coordination. For example, destroying a terrorist training ground may require the US to take action in North Africa or the Middle East. Theoretically, military aid to a country that hosts a terrorist organization creates a disincentive to remove the terrorist problem [13]. In addition, terrorism is a tactical choice for a rebel group when facing a formidable state that the rebels do not want to provoke too much [33].

Both these models suggest that military strength should be a predictor of terrorist attacks. Our algorithm identifies the size of the military as one of the most important predictors of terrorism. As noted in Figure 3, an increase in the size of the military predicts an initial rise in terrorist acts as expected by both the game-theoretic models that predict an increase in the intensity of terrorist attacks, particularly suicide attacks, when targets harden [12]. Nevertheless, further increases in the size of the military keep the risk of terror attacks elevated without increasing terrorist attacks, a potential benefit for a host country receiving military aid. That is, military size increases terrorism at first and then levels off, tracking the prediction from Bapat's model (see Figure 2, p. 311 in [13]).

Equilibrium may also shift from guerrilla warfare to terrorism as a function of the accuracy of a state's military action [33]. If terrorist tactics are more provocative, the probability of a terrorist attack increases. On the flip side, if guerrilla action is more provocative then the probability of guerilla warfare increases. The point is that as the degree of provocation changes there is an equilibrium switch from guerilla warfare to terrorism. Our result in Figure 3 identifies just such an equilibrium switch to increased terrorism as the intensity of guerilla warfare increases. First of all, this means that equilibrium switches to more terrorism are related to guerilla warfare. Thus, our result in Figure 3 supports Carter's model prediction. However, our result also suggests that, if Carter's model is a true reflection of reality, then as guerilla warfare intensifies there is some change in the underlying parameters in a way that makes terrorist action more provocative. Thus, our

results suggest that there may be a relationship between guerilla action and provocation that changes the likelihood of terrorism. This space may bear further theoretical investigation.

Terrorist organizations need to survive to achieve their goals. A strand of the game-theoretic literature is devoted to understanding how terrorist organizations recruit and retain members while overcoming incentive compatibility problems when secrecy is essential.

De Mesquita's game-theoretic model suggests that counterterrorism efforts that reduce economic opportunity can increase terrorist mobilization [39]. In any case, terrorist organizations will put more resources into terrorism (presumably leading to more successful attacks) when they recruit and retain higher-ability terrorists. Therefore, the BDM (2005) model would suggest that, empirically, countries with better economic opportunities would have fewer terrorist attacks. Moreover, he notes that his model suggests, among other things, that ethnically divided societies would see more terrorist attacks and that development aid may reduce terrorism (presumably by increasing economic opportunity).

Our results fail to validate many of the predictions of this model [39]. For example, the variable investment profile includes contract enforcement and risk of expropriation by the state. These variables are components of economic freedom or opportunity. For example, the risk of expropriation reduces the likelihood of economic growth [40]. However, this variable is not an important predictor of terrorism. Moreover, neither ethnic divisions (as measured by the ethnic tensions variable) nor development aid are important predictors of terrorism.

Presumably terrorist organizations mobilize to perpetrate terrorist attacks. However, while factors such as the lack of economic opportunity may indeed affect mobilization, it seems highly unlikely that they affect terrorist attacks. If economic opportunity was important for mobilization it should be able to predict terrorism since terrorism is the purpose for mobilization. This brings into question the role of economic opportunity in explaining mobilization.

Our examples in this section suggest that some game-theoretic models generate validated predictions while others do not. Now all of these models may be causal. Our algorithms make no claims for causality. Yet, if a model is a generalizable explanation of reality, then its predictions should be validated empirically. On this criterion, all models cannot be treated equally. Further, we show how partial dependence plots, by highlighting nonlinear relationships, can help validate game theoretic models that very typically generate hypotheses with nonlinear patterns. Further, these results are data-driven and therefore unbiased by assumptions about any particular theoretical concern. Consequently, empirical results that are consistent with theoretical consequences provide an unbiased validation. Last, once again because our results are data-driven rather than based on theoretical assumptions, they can give us hints about what areas need a theoretical structure. Of course, empirical validation of this new theory should give rise to even more spaces that need theory in an iterative process that slowly erases gaps in knowledge.

7. Conclusions

In this paper, we highlight two aspects of machine learning that can supplement game-theoretic analysis. First, we can sift among competing theoretical models in a theoretically agnostic way to identify those models which have the most predictive salience. A good theoretical model should be able to make predictions. Here, our algorithm suggests that models predicting economic opportunity, development assistance, and ethnic tensions may not be predictively salient. In contrast, those that predict a more formidable target would elicit more terrorist attacks so are predictively salient.

Game-theoretic models, by their very nature, highlight endogenous relationships driven by strategic interactions. Machine learning algorithms, by focusing on predictive accuracy instead of tests of significance, can identify whether a variable is predictive or not even if it is endogenous with the target variable, terrorism. To the extent that causal variables should be predictive, identifying predictive variables can help jumpstart the search for causal links. This process is made more efficient because we can eliminate

variables that are unlikely to be causal because they are not predictive in an empirical framework that is unbiased by endogeneity problems.

Second, game-theoretic approaches often predict nonlinear relationships between variables where equilibriums switch in comparative static scenarios. The partial dependence plots generated by machine learning algorithms can identify these nonlinearities and equilibrium switches in a theoretically agnostic way. Partial dependence plots are, therefore, a particularly suitable testing methodology for game-theoretic comparative statics.

Thus, machine learning techniques can reduce bias and help find better explanations for terrorism. This is important for formulating better counterterrorist policies. These techniques have other benefits as well. For example, they can impute missing data and predictively validate the imputation, and they do not require heroic assumptions about the underlying distribution of data.

Author Contributions: All authors contributed equally in all parts of the project. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: See citations in the text and reference list.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Sandler, T.; Arce, D.G. Terrorism: A game-theoretic approach. *Handb. Def. Econ.* **2007**, *2*, 775–813.
- Wasserstein, R.L.; Lazar, N.A. The ASA Statement on p-Values: Context, Process, and Purpose. *Am. Stat.* **2016**, *70*, 129–133. [[CrossRef](#)]
- Ward, M.; Greenhill, B.D.; Bakke, K.M. The Perils of Policy by p-value: Predicting Civil Conflict. *J. Peace Res.* **2010**, *47*, 363–375. [[CrossRef](#)]
- Basuchoudhary, A.; Bang, J.T.; Sen, T.; David, J. *Identifying the Complex Causes of Civil War: Causal Interpretations of Machine Learning Technologies*; Palgrave-MacMillan: Cham, Switzerland, 2021.
- Shughart, W.F., II. September 11, 2001. *Public Choice* **2002**, *112*, 1–8. [[CrossRef](#)]
- Shughart, W.F., II. An Analytical History of Terrorism, 1945–2000. *Public Choice* **2006**, *128*, 7–39. [[CrossRef](#)]
- Frey, B.S.; Luechinger, S. How to Fight Terrorism: Alternatives to Deterrence. *Def. Peace Econ.* **2003**, *14*, 237–249. [[CrossRef](#)]
- Faria, J.R.; Arce, D.G. Terror Support and Recruitment. *Def. Peace Econ.* **2005**, *16*, 263–273. [[CrossRef](#)]
- Andreozzi, L. Rewarding Policemen Increases Crime. Another Surprising Result from the Inspection Game. *Public Choice* **2004**, *121*, 69–82. [[CrossRef](#)]
- Zhuang, J.; Bier, V.M. Balancing terrorism and natural disasters—Defensive strategy with endogenous attacker effort. *Oper. Res.* **2007**, *55*, 976–991. [[CrossRef](#)]
- Enders, W.; Sandler, T. Distribution of Transnational Terrorism among Countries by Income Class and Geography after 9/11. *Int. Stud. Q.* **2006**, *50*, 367–393. [[CrossRef](#)]
- Berman, E.; Laitin, D.D. Religion, terrorism and public goods: Testing the club model. *J. Public Econ.* **2008**, *92*, 1942–1967. [[CrossRef](#)]
- Bapat, N.A. Transnational terrorism, US military aid, and the incentive to misrepresent. *J. Peace Res.* **2011**, *48*, 303–318.
- Enders, W.; Hoover, G.A. The nonlinear relationship between terrorism and poverty. *Am. Econ. Rev.* **2012**, *102*, 267–272.
- Mullainathan, S.; Spiess, J. Machine learning: An applied econometric approach. *J. Econ. Perspect.* **2017**, *31*, 87–106.
- Hand, D.; Mannila, H.; Smyth, P. *Principles of Data Mining*; MIT Press: Cambridge, MA, USA, 2001.
- Ripley, B.; Venables, W. *R Package 'nnet' Version 7.3-12*; R Foundation for Statistical Computing: Vienna, Austria, 2016. Available online: <https://cran.r-project.org/web/packages/nnet/nnet.pdf> (accessed on 28 June 2021).
- Breiman, L.; Friedman, R.A.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; Chapman and Hall: Boca Raton, FL, USA, 1984.
- Schapire, R.E. The strength of weak learnability. *Mach. Learn.* **1990**, *5*, 197–227. [[CrossRef](#)]
- Freund, Y.; Schapire, R.E. Experiments with a new boosting algorithm. *ICML* **1996**, *96*, 148–156.
- Schapire, R.E. Using output codes to boost multiclass learning problems. *ICML* **1997**, *97*, 313–321.
- Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
- Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140.
- Breiman, L. Random forests; Machine learning. *Mach. Learn.* **2001**, *45*, 5–32.

25. Gassebner, M.; Luechinger, S. Lock, stock, and barrel: A comprehensive assessment of the determinants of terror. *Public Choice* **2011**, *149*, 235–261. [[CrossRef](#)]
26. Banks, A.S.; Wilson, K.A. *Cross-National Time-Series Data Archive*; Databanks International: Jerusalem, Israel, 2015. Available online: <http://www.databanksinternational.com> (accessed on 28 June 2021).
27. Cruz, C.; Keefer, P.; Scartascini, C. *Database of Political Institutions Codebook, 2015 Update (DPI2015)*; Inter-American Development Bank: Washington, DC, USA, 2016.
28. PRS Group. *International Country Risk Guide*; PRS Group: East Syracuse, NY, USA, 2015. Available online: <http://epub.prsgroup.com/products/international-country-risk-guide-icrg> (accessed on 28 June 2021).
29. Solt, F. The Standardized World Income Inequality Database*. *Soc. Sci. Q.* **2016**, *97*, 1267–1281. [[CrossRef](#)]
30. Reynal-Querol, M. Ethnicity, Political Systems, and Civil Wars. *J. Confl. Resolut.* **2002**, *46*, 29–54. [[CrossRef](#)]
31. Mira, A.N. Moscow: Miklukho-maklai ethnological institute at the department of geodesy and cartography of the state geological committee of the soviet union. *USSR* **1964**.
32. Basuchoudhary, A.; Bang, J.T.; Sen, T.; David, J. *Predicting Hotspots: Using Machine Learning to Understand Civil Conflict*; Rowman & Littlefield: Lanham, MD, USA, 2018.
33. Carter, D.B. Provocation and the Strategy of Terrorist and Guerrilla Attacks. *Int. Organ.* **2016**, *70*, 133–173. [[CrossRef](#)]
34. DeRouen, K., Jr.; Sobek, D. State capacity, regime type, and civil war. In *What Do We Know about Civil Wars*; Madon, T.D., Mitchell, S.M., Eds.; Rowman and Littlefield: Lanham, MD, USA, 2016; pp. 59–74.
35. Hendrix, C.S. Measuring state capacity: Theoretical and empirical implications for the study of civil conflict. *J. Peace Res.* **2010**, *47*, 273–285. [[CrossRef](#)]
36. Wilson, M.C.; Piazza, J.A. Autocracies and Terrorism: Conditioning Effects of Authoritarian Regime Type on Terrorist Attacks. *Am. J. Political Sci.* **2013**, *57*, 941–955. [[CrossRef](#)]
37. Varian, H.R. Big data: New tricks for econometrics. *J. Econ. Perspect.* **2014**, *28*, 3–28.
38. Basuchoudhary, A.; Razzolini, L. The evolution of revolution: Is splintering inevitable? *Econ. Peace Secur. J.* **2018**, *13*, 43–54. [[CrossRef](#)]
39. De Mesquita, E.B. The quality of terror. *Am. J. Political Sci.* **2005**, *49*, 515–530. [[CrossRef](#)]
40. Aguiar, M.; Amador, M. Growth in the Shadow of Expropriation. *Q. J. Econ.* **2011**, *126*, 651–697. [[CrossRef](#)]