Toulouse School of Economics

M2 EEE Master Thesis

The Power of the People: Accounting for Public Sentiment in Conflict Prediction

Written By

Luca Poll

Supervised By

Paul Seabright

Augustin Tapsoba

September 24, 2021

Abstract

This paper explores the potential for improved out-of-sample violence forecasting performance through a theoretically informed forecasting model. The model incorporates a set of measures that account for changes in the public sentiment, constructed from the ICEWS event dataset, and forecast violence one quarter ahead on fine-grained Colombian violence data using a Random Forest. Our model performs consistently well in forecasting whether any or two and more violent events occur. It furthermore outperforms the other tested models when predicting the onset of violence and adds to the discussion of the relevance of theoretically grounded models in conflict prediction.

Table of Contents

1 Introduction

Violent conflict leads to mass destruction, human suffering and tragedy. In 2020 alone, violent conflict has caused more than 130,000 fatalities worldwide (ACLED, 2021). Efforts undertaken by the international community to prevent and mitigate conflict through various policy tools are highly dependent on the understanding of the conflict dynamics in the first place.

Scholars have tried for decades to explain the rationale behind violent conflict and tried to identify the causal mechanisms that lead to the outbreak of violence. Many of the causal theories on intrastate conflicts have proven to perform poorly when used for out of sample forecasting and most efforts to create reliable and workable early-warning systems were fruitless. Nevertheless, the conflict prediction literature has made significant progress in forecasting the location of violent conflict. Areas with an increased risk of experiencing violence can be well identified by exploiting time-invariant or slow-moving variables such as measures for the geography or remoteness of a given location, as well as measures of the strength of institutions and the economic performance. The downside of these measures, however, is that the lack of within variation hinders us from forecasting accurately when these areas with an increased risk will eventually experience violent conflict.

We contribute to the existing literature by testing whether a theoretically informed time-varying measure of civilian attitudes enables us to formulate more accurate forecasts with respect to the timing of the violence.[1] Our approach is theoretically grounded in the information-sharing model of Berman, Shapiro, and Felter (2011) and builds on the findings by Bazzi et al. (2021), who identify a set of regressors that allows to predict the location of the violence accurately. We firstly discuss and outline the theoretical mechanism that governs our forecasting model and subsequently test the model empirically on Colombian data, where we try to forecast violence between the government and guerrilla groups. We find that our model performs considerably well when predicting different spec-

[1]Throughout the paper we will refer to violent interactions such as combats, ambushes, harassment, military operations, air strikes, etc between the government and an organized actors within a given country as violence.

ifications of violence one quarter ahead out of sample. The model furthermore outperforms alternative specifications when predicting the onset of violence.

This approach adds value to the literature in that it contributes to the discussion of the relevance of theoretical foundations in conflict forecasting[2] by testing the added value of the inclusion of a theoretically informed measure into a forecasting model. We furthermore propose an innovative way of obtaining the much needed within-variation by relying on ICEWS event data. The finding that civilian attitudes matter in determining the violence outcome is in so far policy-relevant as it supports the argument of the complementarity of counterinsurgency and service provision.[3]

The remainder of the paper is structured as follows: in section 2 we give a brief overview of the relevant literature, in section 3 we present our forecasting model that is then tested in section 4. In section 5 we present the results of the empirical application and section 6 concludes.

2 Related Literature

Violent intrastate conflict has been studied for several decades now and researchers have accordingly developed a multitude of causal approaches as well as forecasting ones in order to improve their predictions and explanations of such events. Prominent causal theories of civil war can be distinguished between those which are grounded in commitment problems and incomplete contracting caused by weak state institutions (Fearon and Laitin, 2003) and those which seek to identify the economic conditions that favor insurgency (Collier and Hoeffler, 2004)[4]. Ward, Greenhill, and Bakke (2010) have shown, however, that those explanatory models perform poorly when used for out of sample forecasting, questioning the relevance of causal theories if these very theories fail to forecast accurately.

Newer approaches, such as the ones from Goldstone et al. (2010), Ward, Metternich, et al.

[2] See Ward, Metternich, et al. (2013), Blair and Sambanis (2020), Beger, Morgan, and Ward (2021), Blair and Sambanis (2021)

[3] For an overview see Berman, Felter, and Shapiro (2018)

[4] For a literature review on the causes of civil war see Blattman and Miguel (2010) or Mueller and Rauh (2018).

(2013) or Hegre, Karlsen, et al. (2013)[5], that seek to evaluate theory through forecasting power, have mostly relied on time-invariant or slow moving structural variables that perform well in predicting the location of violent conflict, but fall short in predicting the timing of violence. Aiming at overcoming this poor prediction of the timing of conflict, different approaches have tried to introduce new variables with sufficient within-variation. These efforts have included the use of cell phone call patterns (Berger, Kalyanaraman, and Linardi, 2014), newspaper text for keyword counts (Chadefaux, 2014) and topic modelling (Mueller and Rauh, 2018), or event data for the use of dyadic event categories and pattern retrieval (Chiba and Gleditsch, 2017; Blair and Sambanis, 2020)[6].

Most of these studies are conducted at the country-year level, however, which faces natural limitations as expounded by Cederman and Weidmann (2017),[7] who instead recommend to focus on forecasts with limited spatial and temporal scope. In one of such studies with a limited spatial and temporal scope[8], Bazzi et al. (2021) apply different machine learning techniques on a large set of regressors disaggregated at the second and third administrative divisions level from Colombia and Indonesia. While they are able to accurately identify locations with high levels of violence, they fall short in predicting the timing of the violence.

This paper relates to the existing literature in different ways: by building on the dataset of covariates in Colombia compiled by Bazzi et al. (2021), it focuses on the study of violence at the disaggregated level as suggested by Cederman and Weidmann (2017). Similarly to the approaches by Chiba and Gleditsch (2017) and Blair and Sambanis (2020), we introduce time-varying variables derived from event data. The underlying theoretical foundation, however, is quite different and to our knowledge there were no attempts to include a measure of public sentiment into violence forecasts. The introduction of these theoretically informed time-varying measures of public sentiment adds value to the existing

[5]For an overview see Hegre, Metternich, et al. (2017).

[6]Other papers used event data to predict future events (Montgomery, Hollenbach, and Ward, 2012) or irregular leadership changes (Beger, Dorff, and Ward, 2018). For a seminal use of event data to conflict prediction see P. A. Schrodt (2006) and P. A. Schrodt (2011).

[7]For a wider discussion of the limits of conflict forecasting see Chadefaux (2017).

[8]Other examples include Blair, Blattman, and Hartman (2017), Witmer et al. (2017) or Hegre, Allansson, et al. (2019).

literature since it not only improves the prediction of the timing of forecasts but also allows to evaluate the information-centric theory by Berman, Shapiro, and Felter (2011). Especially the latter is policy relevant, as the theory sheds light on the importance of the civilians in conflict areas, who can prove to be of high value on the tactical level, both for the government as well as for the rebel side.

3 Model

The inclusion of public sentiment into the violence forecasting model firstly demands to understand how public sentiment can affect the outcome of violence. A good starting point to do so is the information-sharing model proposed by Berman, Shapiro, and Felter (2011), which describes a three-way interaction between the government, the rebels and the public. The aim of this model is not to find a causal explanation for conflict, but to try to identify a common logic at the tactical level where the violence is organized. Building on their findings, we describe how we can make use of changes in the public sentiment in formulating a forecasting model.

3.1 Information-sharing Model

Through the game-theoretic information-sharing model[9], Berman, Shapiro, and Felter (2011) highlight the importance of the private information held by the public and discuss how the decision of sharing this information can be influenced through service provision. This private information is of major importance for the government in the context of an asymmetric conflict when fighting the rebels. For our purposes, we will not focus on the service provision by the government but on the relation between the private information held by the public and the level of violence that is inflicted.

To outline the setting in which the government, the rebels and the public interact, Berman, Shapiro, and Felter (2011, pp.774-778) make several assumptions: as violence is inefficient in the Coasian sense, there must be incomplete contracting between the rebels and the

[9]The information-sharing model is built on a model of criminal street gangs by Akerlof and Yellen (1994). Note that due to the limited space of this thesis we will focus on the intuition of the model instead of the detailed technicalities.

government for violence to occur. Then, the government is assumed to have superior troop mobility and firepower compared to the rebels, can seize territory at any time and can enforce tips at day and night. As a result, the government can remove any threat if the threat can be located.[10] The public is assumed to hold critical information about the rebels, and the government derives a high marginal utility from obtaining such information.

The three players are characterized as follows: the rebels seek social transformation and concessions by inflicting costs on the government through violence; the government seeks to reduce violence through counterinsurgency and service provision; and the public is utility maximizing and can help the government to control a territory by sharing information about the rebels. The public is furthermore assumed to derive utility from services provided by the government or the rebels, holds political preferences, dislikes civilian casualties and fears retaliation.

The interaction of the players itself is structured in the following way: in the first stage, the government and the rebels simultaneously move. The government chooses the level of service provision and the level of counterinsurgency, given its resource constraints.[11] The rebels, on the other hand, choose the level of violence to inflict. After observing the first stage, the public decides on the level of information to share with the government, knowing that this information might be critical in delivering control of the territory to the government.

Berman, Felter, and Shapiro (2018) show that in a setting without uncertainty[12], no violence occurs since the rebels and the government both correctly infer their respective level of support. The rebels will not attempt to ambush the government in areas where they have no support, knowing that they would be given away by the public. The government on the other hand would not station troops in areas without public support, since they know they would be attacked by the rebels and would not obtain intelligence from the

[10]These assumptions describe common features in an asymmetric intrastate conflict.
[11]The government is assumed to not have the capacity to provide services and troops to every location in the country at the same time.
[12]This uncertainty relates to the battle outcome and the civilian attitudes.

public to locate and remove the rebels. With the introduction of uncertainty, however, it is possible to observe violence in equilibrium.[13] In order to deduce this outcome, we shall consider the *perceived probability of winning* by the two competing sides. The rebels and the government both have an unobserved threshold of the perceived probability of winning, above which they are willing to engage in violent confrontation. If this threshold is mutually surpassed, we observe violence. If the threshold is exceeded for only one of the sides, however, we do not observe conflict as not both of the parties are willing to engage in violence.

Hence we can write:

$$y_{it} = \mathbb{1}\{\pi_{1it} \geq \tilde{\pi}_{1it} \cap \pi_{2it} \geq \tilde{\pi}_{2it}\}$$

where y_{it} represents an indicator for whether we observe violence at location i on time t (hence, it can only take values 0 and 1), $\mathbb{1}\{\cdot\}$ an indicator function, π_{1it} the perceived probability of winning for opposing party 1 in location i at time t and $\tilde{\pi}_{1it}$ the threshold above which opposing party 1 is willing to engage in violence in location i at time t. Likewise for party 2. Note that the individual thresholds are unobserved.

3.2 Perceived Probability of Winning and Public Support

After having established that violence occurs in the described setting when both parties mutually perceive a high enough chance of winning a violent confrontation, it is crucial to identify the features influencing this perceived probability of winning. As the term already indicates, reflects the perceived probability a subjective probability which is individually determined and includes case specific factors and assessments. It is most likely, however, that these individual assessments share a common trait when accounting for key determinants.

These key determinants include *inter alia* the military strength, which undoubtedly al-

[13]The uncertainty of the battle outcomes relates to the fact that, despite the asymmetry of force, the government is not guaranteed to win when fighting or the ambushes by the rebels are not guaranteed to be successful when planted. Most importantly, however, civilian attitudes can only be guessed. This is especially hard since attitudes within a village can be highly heterogeneous and hence the government or the rebels can only roughly assume the distribution of attitudes.

ters or diminishes the probability of winning a violent confrontation. In the setting of an asymmetric conflict, however, we indirectly assume that rebels do not have the military strength to win a direct confrontation with the government. Furthermore, we can safely assume that the uncertainty surrounding the estimated military strength of the opponent is limited in a violent conflict with repeated confrontations between both sides, and that the military strength is relatively static or slow-moving across time. Another key determinant of the perceived probability of winning is the level of information that the public shares with the government. If the government has access to this information, it can locate and eliminate threats posed by the rebels. If the government does not have access to the information held by the public, however, the rebels can freely operate from the underground and successfully ambush the government as there is no clearly defined front. Since these dynamics are known to the public when deciding on the level of information to share, they can actively weaken or strengthen either side via their support. Therefore, the public support constitutes a key determinant in forming the perceived probability of winning. Ultimately, we can identify location/time idiosyncratic factors that describe battle-related uncertainty to influence the perceived probability of winning but can assume these effects to be zero on average.

3.3 Including Public Sentiment as a Time-varying Regressor

Adding to the previous finding that the perceived probability of winning itself is determined *inter alia* by the estimated military strength and the assumed public support for the competing sides, we get closer to deriving a testable hypothesis. Before doing so, however, it is important to discuss the mechanisms at work.

In their paper Bazzi et al. (2021) identify as most important predictors of violence measures of remoteness like the distance to major cities and road access, geographic traits like terrain ruggedness or measures of the economic structure like sectoral shares, agricultural features or mineral presence. All these predictors are time-invariant and might jointly explain the presence of a commitment problem through weak institutions (Fearon and Laitin, 2003) or describe economic conditions that allow the rebels to raise revenues or facilitate recruitment (Collier and Hoeffler, 2004). In the understanding of the proposed model, however, these structural predictors are not considered at the causal level, but at the tactical level instead.[14] As such, rugged terrain or no road access in an area might increase the perceived probability of winning for the rebels since the relative military strength might be positively impacted due to the terrain favoring guerrilla tactics. Analogously, rural and agriculturally based communities might align closer with the political ideologies of rebel groups than with the ideologies of the government and hence naturally lower the level of information shared with the government. Since these factors are time-invariant, however, they can be useful in identifying contested areas, but are of little use when trying to predict the timing of violence.

Focusing solely on changes in the public support might prove more fruitful in obtaining the much needed within variation. We can think of public support as being the result of continuous interactions between the government, the rebels and the public, like in the case of protests, arrests, concessions, and scandals. These interactions are closely linked to the public sentiment which determines the public support and can be considered as volatile. As a result, we can model the perceived probability of winning as being influenced by structural, time invariant factors or time varying factors like public support:

$$\pi_{pit} = f\left(V_{it}, X_i\right)$$

where the perceived probability of winning of party p in location i at time t is a function of a set of time-varying factors V_{it} and a set of time-invariant structural factors X_i. As discussed before, the public support is a key determinant of the time varying factors, which in turn can be expressed as a rigid outcome of accumulated changes in public sentiment

[14]The tactical level refers to the tactical decisions that are taken by the opposing groups on the field when both groups already engaged in violence.

and other observed and unobserved factors:

$$sup_{pit} = g\left(\Delta sent_{(p,q)\,i,\,t-1},\ Z_{i,\,t-1}\right)$$

where sup_{pit} describes the public support for party p in location i at time t which can be modeled as a function of a set of past changes in the public sentiment between party p and public q up to time $t-1$ and Z_{it}, which describes other factors influencing the public support. Changes in the public sentiment might shift the public support from one to the other side and push the perceived probability of winning above or below the respective thresholds if close enough to the tipping point. As a result, changes in the public sentiment can indicate a regime change from no violence to violence if the public support is ambiguous enough. Anticipating accurately this shift would allow us to formulate more precise forecasts on the timing of conflict. By leveraging the volatile nature of public sentiment, we can express the forecasting model as follows:

$$\mathbb{E}\left[y_{i,\,t+1}\right] = h\left(\Delta sent_{(p,q)\,i,\,t},\ Z_{it},\ X_i\right)$$

where the expected violence outcome $\mathbb{E}\left[y_{i,\,t+1}\right]$ can be described by a set of different measures $\Delta sent$ capturing the current and past changes in public sentiment between p and q at location i up to time t. Z_{it} describes a set of other variables influencing the public support of which some can be observed and others not. X_i describes the set of time-invariant regressors influencing the perceived probability of winning.

4 Empirical Application: Guerrilla Groups in Colombia

The proposed model is tested empirically on fine-grained high quality data from Colombia quarterly from 2001 until 2014 with the underlying hypothesis that the inclusion of a measure of the public sentiment improves the predictability of the timing of violence. As the setting of the model is not coherent with all types of intrastate violence, we firstly want to discuss to what extend the features of the violence in Colombia fit with our model's assumptions.

4.1 Violence in Colombia

Ever since the assassination of the liberal political leader Jorge Gaitán, which sparked a ten year long civil war between the conservative and the liberal party in 1948, has the country experienced high levels of violence and destruction. According to the Observatorio de Memoria y Conflicto (2021) did the violent conflict claim more than 267'000 lives between 1958 and 2021 and displaced over five million persons (Grupo de Memoria Histórica, 2016).

Most of these victims can be traced back to the long-running conflict between left-wing guerrilla groups, the government and right-wing paramilitary groups. The largest guerrilla group[15], the Revolutionary Armed Forces of Colombia (FARC), was founded in 1964 as a Marxist-Leninist guerrilla group to represent and defend the rural population's interest and originally aimed at overthrowing the government (Center for International Security and Cooperation Stanford University, 2019). The troop strength of the guerrilla groups as well as the level of violence caused by the guerrillas has varied over the years and has been steadily decreasing since its peaks in 1999. This decline started after the "no más" protests, where 13 million Colombians marched on the streets calling for immediate ceasefires and is likely to be caused by a 9 billion USD military aid program by the US government to support the Colombian government in fighting the drug trade in 2000, combined with a stringent anti-guerrilla policy of president Àlvaro Uribe (2002-2010). President Uribes successor, Juan Manuel Santos, brought the FARC back to the negotiation table in 2012 and, after a period of mutual violations of agreed ceasefires, paved the way for the 2016 peace deal, which ended the Colombian civil war officially. Yet, the violence continues, since not all guerrilla groups have demobilized and although the FARC has officially turned their struggle into a political one, many former FARC fighter regrouped in FARC dissident groups, also known as the ex-FARC Mafia.

[15] Other guerrilla groups include the National Liberation Army (ELN), the Popular Liberation Army (EPL) or the 19th of April Movement (M-19), of which the latter two demobilized in the late 80's and early 90's.

The type of conflict that is related to the guerrilla's revolutionary struggle can be identified as an intrastate succession conflict, that is asymmetric and information-centric. Even though the FARC had an estimated troop strength of around 16'000 fighters and the ELN additional estimated 3'500 troops in 2001 (Center for International Security and Cooperation Stanford University, 2019), it's military capacities were at any time outmatched by the governmental forces, especially after receiving the US military support. The use of hit and run guerrilla tactics demanded to have refuge areas with civilian sympathizers as the information about the location of the guerrilla camps was crucial for the government to fight the guerrillas. As of such, the conflict can be considered information-centric.

4.2 Data

The data is made up of three different sources: the data on violent events ($y_{i,\,t+1}$) is taken from Observatorio de Memoria y Conflicto (2021)[16], the structural variables (X_i) are taken from the Bazzi et al. (2021) dataset for Colombia, and the measures for change in public sentiment up to the previous period ($\Delta sent_{it}$) are created from the ICEWS event dataset (Boschee, Lautenschlager, O'Brien, et al., 2015). For more details on the respective datasets see Appendix A. The data is aggregated quarterly at the municipality level[17], and the study is limited to rural municipalities, since the dynamics of conflict are arguably different in rural compared to urban areas.[18]

The violence outcome, y_{it}, is coded in the baseline specification as a binary variable indicating whether the municipality i experienced a violent event at quarter t or not. Violent events include combats, ambushes, military operations, attacks on military bases, harassment and air attacks as indicated by the Observatorio de Memoria y Conflicto (2021). During our period of interest at least 4% and up to 30% of the municipalities were engaged into violence at every point in time, as we can see from Figure 1. The municipalities

[16] The high level of detail of the violence data allows us to also construct variables describing other factors that influence the public support (Z_{it}).

[17] The municipios represent the second-level administrative divisions in Colombia. We follow the approach by Dube and Vargas (2013) in using the 1988 municipality border as several additional municipalities were created since the new constitution in 1991.

[18] We define rural areas through the population density. Municipalities with an average population density one standard deviation above the sample mean population density are excluded. Through this measure, 26 municipalities with the highest population densities in Colombia are not considered in the study.

were not equally affected by the violence, however, as we can see in Figure 2. The most violent municipalities had up to 376 violent events across the 56 quarters of the study while others have not experienced a singe event. Municipalities that experienced a large number of violent events were not violent throughout all periods, as we can see from Figure 3. The majority of municipalities experienced violence in less than 25% of the quarters under study.

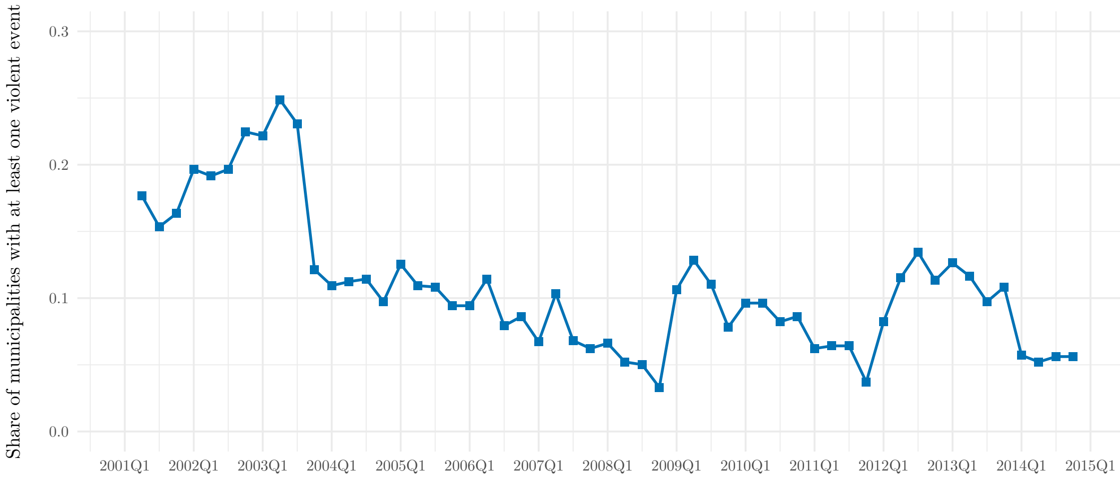


Figure 1: The time series shows the share of municipalities that experienced violence across the quarters of the study.

The measures for changes in the public sentiment take as a starting point the ICEWS dataset. Each of the registered events is associated to a CAMEO event code (P. Schrodt, 2012) which in turn is associated with a CAMEO intensity score, indicating the intensity of an event. The scores range from -10 (very negative sentiment) to +10 (very positive sentiment), where "engage in unconventional mass violence" for instance would be associated with an intensity of -10 and "retreat or surrender military" would have an intensity score of +10 while "engage in diplomatic cooperation" would be considered with an intensity score of +3.5 and "investigate" with -2. We derive different measures for the change in public sentiment from the ICEWS event data. One measure is constructed by aggregating the intensity scores quarterly for each municipality and each relationship between the actors.[19] We include the first three lags to control for longer delays besides the measure

[19] The relationships respect the direction of the interaction and can be public to government or rebels to public for instance.

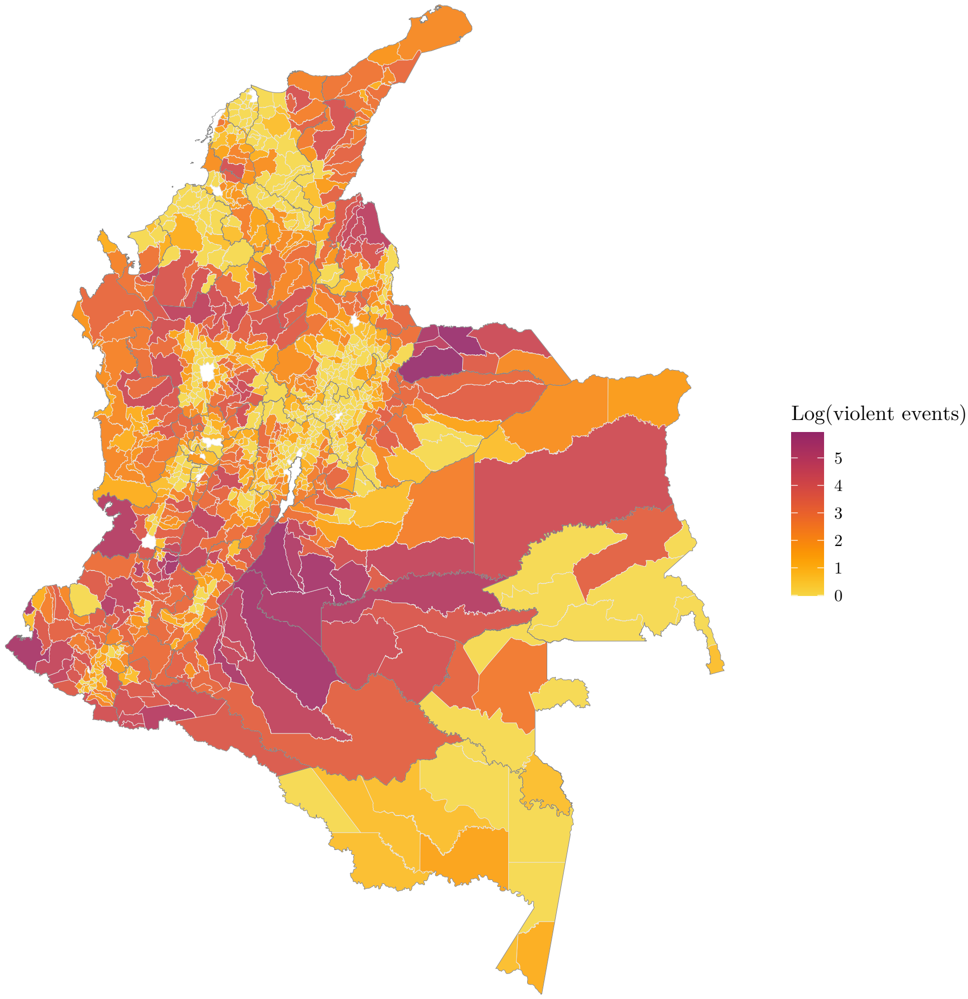


Figure 2: Map of the total count of violent events across the period of interest by municipality. Note that the numbers are reported in logs only for visualization purposes.

of quarter t that is used to forecast violence in quarter $t+1$. In the same fashion, we include a measure that considers the average sentiment instead of the aggregated sentiment with the respective first three lags. We furthermore define a less specific measure by disregarding the direction of interaction between an actor pair when aggregating or averaging the intensity scores at the quarterly level by location and call it net sentiment. Just like for the other measures, we include the first three lags for the aggregated and the first three lags for the average net sentiment. To account for longer-term public sentiment, we include the cumulative aggregate event intensity for every municipality and actor rela-

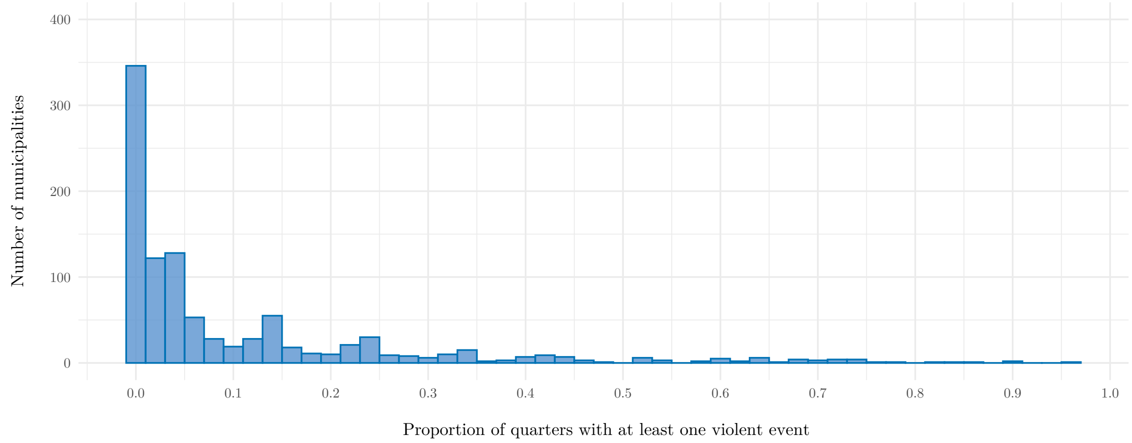


Figure 3: Histogram of the proportion of quarters in which a municipality experienced violence over all quarters of the study.

tionship (respecting the direction) as well as the net sentiments from $t = 1$ up to t. As we can see in Figure 5, does the sum of intensity codes vary considerably across the quarters for all of the directional actor pairs. We can see furthermore that the sum is in most cases negative, which might relate to the fact that negative events are likelier to be reported than positive ones.[20]. The extend to how negative these interactions are, however, varies by relationship as we can see in Figure 5.

Since the public support is also defined by factors other than the change in public sentiment, we make use of the high level of detail of the Observatorio de Memoria y Conflicto (2021) violence data to create regressors that enter into Z_{it}. As mentioned in section 3, does the public dislike civilian casualties, which is why we expect that the public support diminishes if civilians get killed or injured. Therefore, we include the aggregate of injured civilians and civilian victims by quarter and municipality with the respective first three lags. We furthermore make use of the information of who initiated the violence. The rationale is that the party that took the initiative to attack had to be confident enough to have a sufficiently high perceived probability of winning which is made up through the military strength and the public support. As of such, the quarterly count of initiatives by party and location can be considered as an indication for past public support. As for the other variables, we include the respective first three lags.

[20]For a discussion about the media bias see Chadefaux (2017).

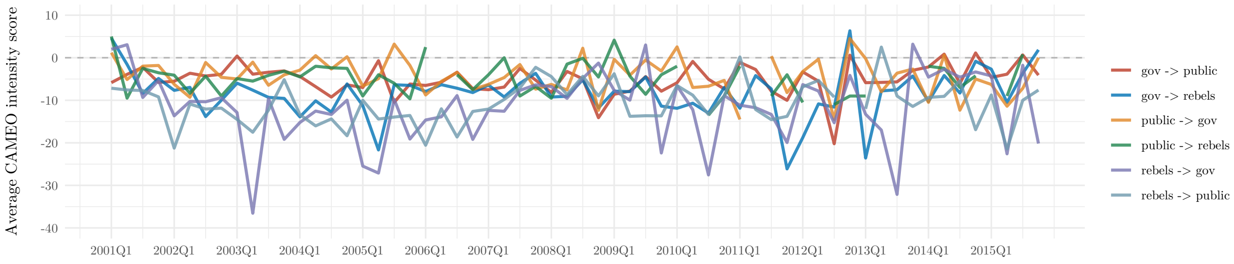


Figure 4: Time series of the average quarterly CAMEO intensity score sum across the municipalities in the sample by actor pair and direction of the interaction

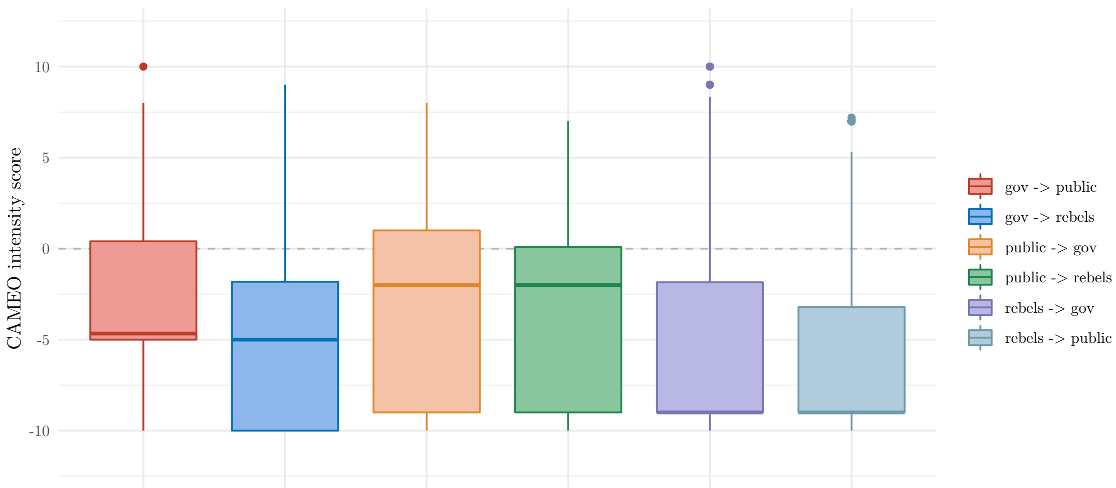


Figure 5: Boxplots of the CAMEO intensity scores by actor pair and direction of the interaction

The structural variables X_i that are taken from Bazzi et al. (2021) contain various regressors describing inter alia the population, the remoteness, the geography, the history or the production capacities of a municipality. Most of these regressors are time-invariant, while some are slow-moving. To rely only on the structural information contained in these variables, we simplify them to the average level between 1991 and 2013.

4.3 Empirical Strategy

We make use of a Random Forest that is used to formulate our forecasts. This supervised machine learning algorithm is not only more accurate in predicting conflict outcomes than the most commonly used logit or probit models (Muchlinski et al., 2016), but spares us furthermore from making parametric assumptions due to the nonparametric nature. The

Random Forest algorithm is explained in more detail in Appendix B.1. Random Forests are generally considered as black boxes, which hinders us from making causal inference. If the proposed model delivers enhanced forecasts, however, it indicates that the model can be a step into the right direction and further causal inference should be conducted. To estimate and evaluate our model we undertake the following steps:

1. We take a subset of the data from Q1 2001 to Q4 2010 and use five-fold Cross Validation for hyperparameter tuning. We previously chose the number of trees to grow to be 250[21] and tune the optimal number of variables that should be considered at each split (P).

2. For each t between Q1 2011 and Q3 2015 we split the sample into a training set$_t$ (all observations up to t) and a test set$_t$ (observations in t including the violence outcome from $t+1$).

   (a) We fit a Random Forest to the training set$_t$ using the optimal P obtained in step 1 and downsample observations that had no violence to account for the imbalanced outcome.

   (b) We use the fitted model to predict the violence outcome in $t+1$ from the test set$_t$.

   (c) We compute the area under the Receiver Operating Characteristics (ROC) curve (AUC$_t$).

3. The average AUC is computed from the individual AUC$_t$ and reported as the measure of evaluation.

This empirical strategy allows us to use all available information at every t to predict the violence outcome in $t+1$.

[21] The number of trees only needs to be lage enough. We chose to grow 250 trees due to computation power constraints. The Out-Of-Bag error rate stabilized for most of the tested specifications at around 75 trees already, which is why 250 trees can be considered as sufficient.

5 Results

We follow this procedure for different model specifications and different dependent variables. Our main focus lies on predicting whether a municipality will experience conflict one quarter ahead, which is why *any violent event* is chosen as the first dependent variable. As a second dependent variable, we define *violence hotspot* as municipalities that experience two or more violent events in one quarter and try to forecast such hotspots one quarter before. The third dependent variable, *violence onset*, captures whether a municipality experiences violence in a quarter after not having experienced any violent event during the four previous quarters. We test our model against two baseline specifications, where the first specification relies simply on the first four lags of the dependent variable and the second specification uses the 25 most relevant and time-invariant structural variables from the Bazzi et al. (2021) dataset. Our proposed model is additionally compared to a specification that includes the structural variables as well as a set of the detailed violence history for the past four quarters respectively. The results are presented in Table 1.

Table 1: Results of the one-quarter ahead out of sample forecast using a Random Forest

	Autoregressive	Structural	Structural + Public Support	Structural + Violence History
	(1)	(2)	**(3)**	(4)
Any violent event	0.855	0.919	0.939	0.939
≥ 2 violent events	0.903	0.935	0.952	0.953
Violence onset	-	0.857	0.875	0.827

Notes: The reported AUC is computed as the average AUC obtained from the out of sample prediction as outlined in section 4.3. *Violence onset* is coded as an indicator that is equal to one for the first quarter experiencing violence after four consecutive quarters of no violent event. The *Autoregressive* model uses as potential regressors the first four lags of whether any violent event occurred, allowing for an autoregressive process op to $AR(4)$. Note that it is not possible to run the AR specification for the *violence onset* as municipalities that experience an onset by definition have not experienced violence in the precedent quarters. The *Structural* model includes the 25 most relevant structural variables from the Bazzi et al. (2021) dataset. The *Structural + Public Support* model includes additional to the structural variables the regressors for public support as explained in section 4.2. The *Violence History* model contains detailed information about past violence like the number of victims, injured and captured on both sides as well as on the civilian side, the lagged number of violent events, a count of which side that had the military advantage as well as a count of who took the initiative to attack and a count on the respective type of violence as defined by Observatorio de Memoria y Conflicto (2021).

We shall first consider the predictive performance of the different specifications when forecasting whether a municipality experiences conflict or not. The measure of evaluation, the average AUC[22], indicates hereby how well the model distinguishes between positive (violence) and negative (no violence) cases. We see that relying solely on the lags of the dependent variable performs comparably poor. Using the time-invariant regressors in the structural specification already allows to predict violence considerably better with an average AUC of 0.919. This performance is outmatched through our proposed model, however, as it yields an average AUC of 0.939 which indicates that the changes in public sentiment indeed contain information that can be leveraged to improve the forecasts. The fourth specification yields the same average AUC as our proposed model. The result of this less theoretically grounded specification suggests that the detailed past violence is equally valuable in predicting future violence as the public support model. When looking at the respective prediction performances for the violence hotspots we see the exact same pattern across the different specifications, although the autoregressive model performs comparably better now. This might be due to the fact that violence hotspots experience violence not only more often during a quarter but also across more quarters as shown in Figure 2 and Figure 3 which hence makes it easier to predict violence where it has occurred before. The prediction performance for the violence onset dependent variable is generally lower across the specifications than for the other two dependent variables. This reflects the general difficulty in predicting new violence onsets. Also for this dependent variable the proposed model performs considerably better than the time-invariant structural predictors and on top outperforms the specification that relies on the detailed past violence. The poor performance can be explained by the fact that municipalities that experience an onset of violence do not have data on recent past violence since no violence has occurred.

In summary we can conclude that our proposed model performed consistently well throughout the different dependent variables. In all instances it outperformed the time-invariant model and the simple autoregressive model. It furthermore performed better than the model including the detailed violence history when predicting violence onset which is especially valuable since most forecasting models fail to predict violence outbreaks where it has not occurred before.

[22]For a more detailed overview of the ROC AUC see Appendix A.

6 Discussion

Throughout the course of this paper we assessed whether a theoretically informed measure of public support can improve the prediction of violent events. We compare our method against three alternative forecasting models based on lags of observed violence, on only time-invariant structural variables and on time-invariant structural variables paired with a detailed violence history.

We find that the public support model performs consistently well across different dependent variables and outperforms the other specifications when forecasting the onset of violence. When forecasting whether any violent event occurs one quarter ahead, the public support model and the specification relying on the detailed past violence perform similarly[23]. This holds also true when predicting violence hotspots which we define as municipalities with two or more violent events in a given quarter. We suspect the outlined public support model to be more potent, however, due to data quality issues related to the measure of the public sentiment that is obtained from the ICEWS event dataset. We observe ICEWS events mostly in the more populated areas since the events are automatically coded from news text which is generally more accessible in populated areas. The Colombian guerrilla groups rely on support of the population also in the most rural areas, however, which is why we expect a study that measures public sentiment through survey data to leverage even more information contained in the change in public sentiment that could be used for prediction.

The results obtained from the empirical application support the relevance and the potential of civilian attitudes when used for out of sample prediction nevertheless. A good starting point for trying to understand the underlying causal mechanism in detail would be to extend the information-sharing model by Berman, Shapiro, and Felter (2011) formally and test it empirically on another country.

[23]Both specifications reach an average Area Under the Receiver Operating Characteristic Curve (ROC AUC) of 0.939.

References

ACLED (2021). *Armed Conflict Location & Event Data Project (ACLED)*. Accessed: 23-09-2021. URL: %7Bhttps://acleddata.com/dashboard/#/dashboard%7D.

Akerlof, G. and J. Yellen (1994). "Gang Behavior, Law Enforcement, and Community Values". In: *Values and Public Policy*. Washington: Brookings Inst., p. 26.

Bazzi, S., O. Dube, R. Blair, and C. Blattman (2021). "The Promise and Pitfalls of Conflict Prediction: Evidence from Colombia and Indonesia". In: *Review of Economics and Statistics*. DOI: doi.org/10.1162/rest_a_01016.

Beger, A., C. Dorff, and M. Ward (2018). "Irregular leadership changes in 2014: Forecasts using ensemble, split-population duration models". In: DOI: 10.1016/j.ijforecast.2015.01.009.

Beger, A., R. Morgan, and M. Ward (2021). "Reassessing the Role of Theory and Machine Learning in Forecasting Civil Conflict". In: *Journal of Conflict Resolution*, pp. XXX–XXX.

Berger, D., S. Kalyanaraman, and S. Linardi (2014). "Violence and cell phone communication: Behavior and prediction in Coˆte d'Ivoire". In: *Unpublished*. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2526336.

Berman, E., J. Felter, and J. Shapiro (2018). *Small wars, big data: The information revolution in modern conflict."* New Jersey: Princeton University Press.

Berman, E., J. Shapiro, and J. Felter (2011). "Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq". In: *Journal of Political Economy* 119.4. Publisher: The University of Chicago Press, pp. 766–819. DOI: 10.1086/661983.

Blair, R., C. Blattman, and A. Hartman (2017). "Predicting local violence: Evidence from a panel survey in Liberia". In: *Journal of Peace Research* 54.2, pp. 298–312. DOI: 10.1177/0022343316684009.

Blair, R. and N. Sambanis (2020). "Forecasting Civil Wars: Theory and Structure in an Age of "Big Data" and Machine Learning". In: *Journal of Conflict Resolution* 64.10, pp. 1885–1915. DOI: 10.1177/0022002720918923.

Blair, R. and N. Sambanis (2021). *Is Theory Useful for Conflict Prediction? A Response to Beger, Morgan, and Ward**. Tech. rep.

Blattman, C. and E. Miguel (2010). "Civil War". In: *Journal of Economic Literature* 48.1, pp. 3–57. DOI: 10.1257/jel.48.1.3.

Boschee, E., J. Lautenschlager, S. O'Brian, S. Shellman, J. Starz, and M. Ward (2015a). "ICEWS Coded Event Data ReadMe". In: Harvard Dataverse. DOI: https://doi.org/10.7910/DVN/28075/46OVOW.

– (2015b). "ICEWS Events and Aggregations". In: Harvard Dataverse. DOI: https://doi.org/10.7910/DVN/28075/46OVOW.

Boschee, E., J. Lautenschlager, S. O'Brien, S. Shellman, J. Starz, and M. Ward (2015). *ICEWS Coded Event Data*. Version V30. DOI: 10.7910/DVN/28075.

Cederman, L.-E. and N. Weidmann (2017). "Predicting armed conflict: Time to adjust our expectations?" In: *Science* 355, pp. 474–476. DOI: 10.1126/science.aal4483.

Center for International Security and Cooperation Stanford University (2019). *Mapping Militant Organizations. "Revolutionary Armed Forces of Colombia"*. Accessed: 20-09-2021. URL: https://cisac.fsi.stanford.edu/mappingmilitants/profiles/revolutionary-armed-forces-colombia-farc#text_block_17686.

Chadefaux, T. (2014). "Early warning signals for war in the news". In: *Journal of Peace Research* 51.1, pp. 5–18. DOI: 10.1177/0022343313507302.

– (2017). "Conflict forecasting and its limits". In: *Data Science* 1.1-2, pp. 7–17. DOI: 10.3233/ds-170002.

Chiba, D. and K. Gleditsch (2017). "The shape of things to come? Expanding the inequality and grievance model for civil war forecasts with event data". In: *Journal of Peace Research* 54.2. Publisher: SAGE Publications Ltd, pp. 275–297. DOI: 10.1177/0022343316684192.

Collier, P. and A. Hoeffler (2004). "Greed and Grievance in Civil War". In: *Oxford Economic Papers* 56.4, pp. 563–595. URL: https://www.jstor.org/stable/3488799.

Dube, O. and J. Vargas (2013). “Commodity Price Shocks and Civil Conflict: Evidence from Colombia”. In: *The Review of Economic Studies* 80.4 (285), pp. 1384–1421. ISSN: 0034-6527.

Fearon, J. and D. Laitin (2003). “Ethnicity, Insurgency, and Civil War”. In: *The American Political Science Review* 97.1, pp. 75–90. URL: https://www.jstor.org/stable/3118222.

Goldstone, J. A., R. H. Bates, D. L. Epstein, T. R. Gurr, M. Lustik, M. Marshall, J. Ulfelder, and M. Woodward (2010). “A Global Model for Forecasting Political Instability”. In: *American Journal of Political Science* 54.1, pp. 190–208. DOI: 10.1111/j.1540-5907.2009.00426.x.

Grupo de Memoria Histórica (2016). *Basta Ya! Colombia: Memories of War and Dignity*. Tech. rep. Bogotá: CNMH.

Hegre, H., N. Metternich, H. M. Nygard, and J. Wucherpfennig (Mar. 2017). “Introduction: Forecasting in peace research”. In: *Journal of Peace Research* 54.2, pp. 113–124. DOI: 10.1177/0022343317691330.

Hegre, H., M. Allansson, et al. (Mar. 2019). “ViEWS: A political violence early-warning system”. In: *Journal of Peace Research* 56.2, pp. 155–174. DOI: 10.1177/0022343319823860.

Hegre, H., J. Karlsen, H. Nygård, H. Strand, and H. Urdal (2013). “Predicting Armed Conflict, 2010-2050”. en. In: *International Studies Quarterly* 57.2, pp. 250–270. DOI: 10.1111/isqu.12007.

Lockheed Martin (July 2021). *ICEWS*. URL: https://www.lockheedmartin.com/en-us/capabilities/research-labs/advanced-technology-labs/icews.html.

Louppe, G. (2014). “Understanding Random Forests”. PhD Dissertation. University of Liège.

Montgomery, J. M., F. M. Hollenbach, and M. Ward (2012). “Improving Predictions using Ensemble Bayesian Model Averaging”. en. In: *Political Analysis* 20.3, pp. 271–291. DOI: 10.1093/pan/mps002.

Muchlinski, D., D. Siroky, J. He, and M. Kocher (2016). "Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data". In: *Political Analysis* 24.1, pp. 87–103. DOI: 10.1093/pan/mpv024.

Mueller, H. and C. Rauh (2018). "Reading between the lines: Prediction of political violence using newspaper text". In: *American Political Science Review* 112.2, pp. 358–375. DOI: 10.1017/S0003055417000570.

Observatorio de Memoria y Conflicto (2021). *Sistema de Información de Eventos de Violencia del Conflicto Armado Colombiano*. Accessed: 2021-09-20. URL: %7Bhttp://micrositios.centrodememoriahistorica.gov.co/observatorio/portal-de-datos/el-conflicto-en-cifras/%7D.

Schrodt, P. (2012). *CAMEO - Conflict and Mediation Event Observations - Event and Actor Codebook*. Tech. rep. Version 1.1b3. Department of Political Science: Pennsylvania State University. URL: http://eventdata.psu.edu/.

Schrodt, P. A. (2011). "Forecasting Political Conflict in Asia using Latent Dirichlet Allocation Models". en. In: *Paper presented at the European Political Science Association Annual Meeting*. Dublin, p. 30. URL: https://parusanalytics.com/eventdata/papers.dir/Schrodt.EPSA.2011.final.pdf (visited on 09/18/2021).

– (2006). "Forecasting Conflict in the Balkans using Hidden Markov Models". en. In: *Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention*. Ed. by R. Trappl. Advances in Group Decision and Negotiation. Springer Netherlands, pp. 161–184. DOI: 10.1007/1-4020-4390-2_8.

Ward, M., A. Beger, J. Cutler, M. Dickenson, C. Dorff, and B. Radford (2013). "Comparing GDELT and ICEWS event data". In: *Analysis* 21, pp. 267–297.

Ward, M., B. Greenhill, and K. Bakke (2010). "The perils of policy by p-value: Predicting civil conflicts". In: *Journal of Peace Research* 47.4, pp. 363–375. DOI: 10.1177/0022343309356491.

Ward, M., N. Metternich, C. L. Dorff, M. Gallop, F. Hollenbach, A. Schultz, and S. Weschle (2013). "Learning from the Past and Stepping into the Future: Toward a New Generation of Conflict Prediction". In: *International Studies Review* 15.4, pp. 473–490. DOI: 10.1111/misr.12072.

Witmer, F., A. Linke, J. O'Loughlin, A. Gettelman, and A. Laing (2017). "Subnational violent conflict forecasts for sub-Saharan Africa, 2015–65, using climate-sensitive models". In: *Journal of Peace Research* 54.2, pp. 175–192. DOI: 10.1177/0022343316682064.

A Data Appendix

A.1 OMC dataset

The data on violent events and the violence history is obtained from the 'military actions' dataset extracted from the Observatorio de Memoria y Conflicto (2021) database. The database was created in cooperation with the *Centro Nacional de Memoria Hístorica* (CNMH) and can be considered the continuation of the *Basta Ya!* initiative.

The dataset reports violent events and the respective characteristics such as the time, the location, the actors involved, the type of violence, the side that took the initiative to engage in violence and the side that obtained the military advantage[24] among others. It also includes a detailed and disaggregated[25] count of the victims, the injured and the captured and spans from 1958 to 2021.

The data entries are complied out of 611 different sources[26] which include national archives, NGOs, church organizations, victims associations, news reports and many more. Each data entry is individually checked and included into the dataset if it meets the OMC's standards. In Figure 6 we see that the total number of violent events that is included in the dataset varies considerably throughout he sample. This is mostly due to the number of combats which mark three periods of high violence levels.

A.2 ICEWS dataset

The Integrated Crisis Early Warning System (ICEWS) event data records interactions between socio-political actors, which can be cooperative or hostile interactions between individuals, groups, sectors and nation states. The dataset spans from the 1st of January 1995 to now and is issued on a weekly basis. It includes a total of more than 19 million events[27] that are automatically coded from newspaper text using the BBN ACCENT event

[24]The military advantage is determined by comparing the number of victims and injured on both sides and furthermore considering the damage done to military equipment, the theft of war material or the hostages captured

[25]Disaggregated into combatants of the respective group involved and civilians

[26]As of September 2021

[27]As of 21.07.2021

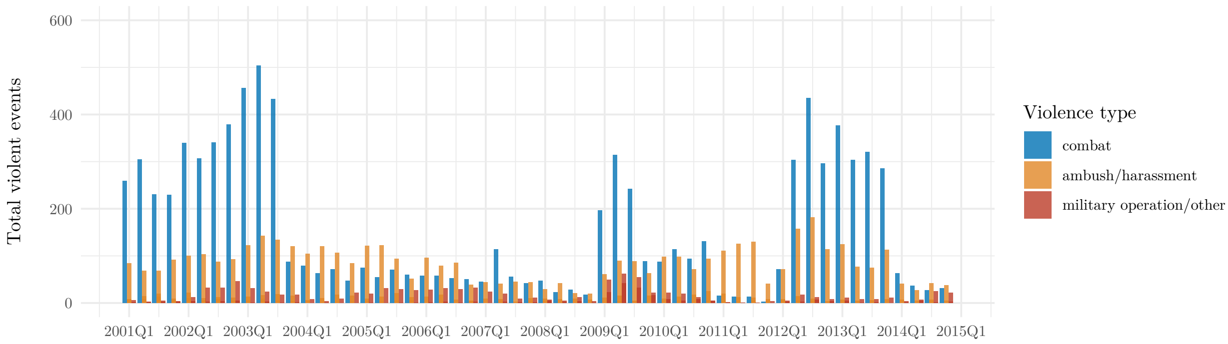


Figure 6: Bar chart of the total count of violent events across the municipalities by type of violence.

coder and the text corpus is built from English, Spanish, Portuguese and Arabic news stories. (Boschee, Lautenschlager, O'Brian, et al., 2015a; Lockheed Martin, 2021)

The ICEWS event dataset is furthermore considered to be the most influential and most thoroughly studied event dataset. It tends to miss some events in comparison to the GDELT event dataset[28], while the GDELT dataset, however, is overstating the number of events substantially (Ward, Beger, et al., 2013). The events are reported as triples: the source, the event type and the target. The event types are categorized via the CAMEO event scale, while the source and the target are identified as actors and agents with the corresponding sectors through the ICEWS dictionaries.

Actors, Agents & Sectors: The source and target represent the "who" part of the event and are presented as either actors or agents. The agents reflect individuals or groups that are common to most of the countries (like police, protestors, students, etc.) while the actors are either 'named actors' that are formally defined in the ICEWS Actor Dictionary or 'composite actors' which are not individually listed but are created by combining a listed actor with an agent or a sector. The sectors are "general affiliations which can be applied to an actor via its defined memberships" and are divided into five hierarchical levels (Boschee, Lautenschlager, O'Brian, et al., 2015b, p.7). The top-level differentiates simply between 'national sectors' and 'unaffiliated sectors', while the second level distinguishes into more fine-grained sectors like *government*, *parties*, *dissident*, *social*, *nongovernmental*

[28] Another widely used event dataset

organizations/ activists, *elite* or *unidentified forces* for the national sector.

Event type: The events on the other hand are classified following the CAMEO Codebook (P. Schrodt, 2012) and hence consist of 292 distinct event types which relate to 20 different categories. Besides the respective CAMEO code are also the event intensities reported, which are expressed though the CAMEO intensity scores that are assigned to the individual CAMEO codes by the CAMEO scale. These values represent the hostility or the cooperation implied by a respective event type ranging from -10 (hostile) to 10 (cooperative). (Boschee, Lautenschlager, O'Brian, et al., 2015b)

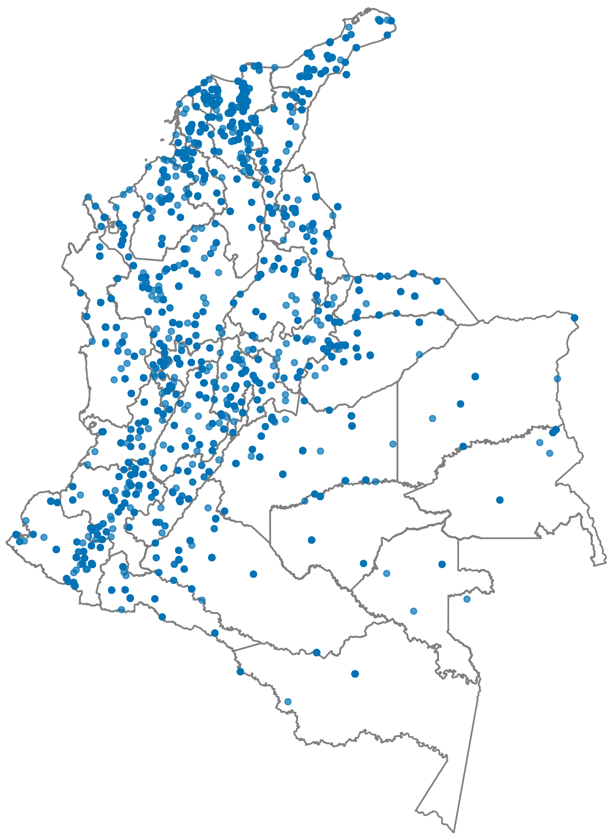


Figure 7: Map of Colombia with the locations of the events extracted from ICEWS.

We filter the ICEWS database for the information that we need for the inclusion into the forecasting model. These filter steps include to just extract domestic events where the source and the target sector are not the same. We exclude furthermore events in the municipalities with a high population density. As we can see in Figure 7 do we observe events mostly in the more populated areas. Once extracted, we assign the sectors/actors to the categories gov, rebels or public which leaves us with interactions as shown in Figure 8.

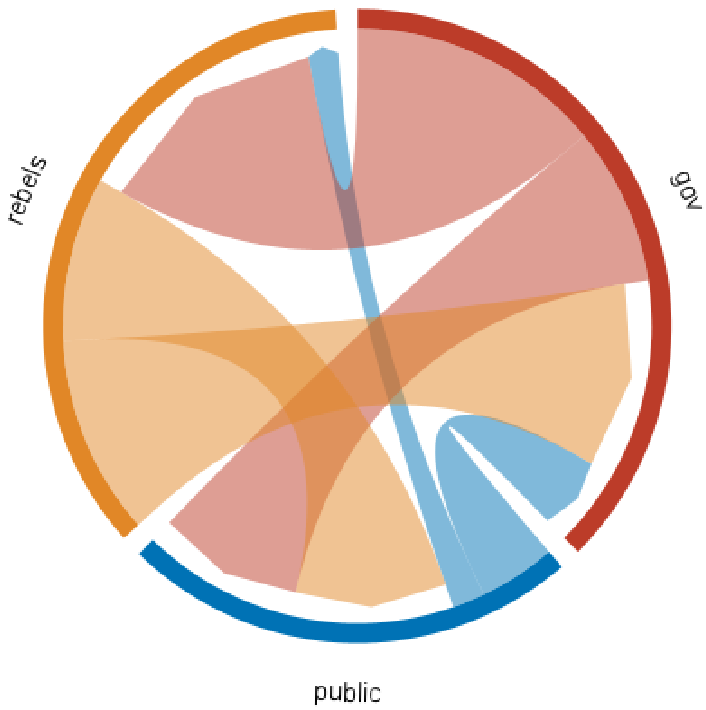


Figure 8: The chord diagram shows the direction and quantity of actor interactions in the extracted sample. The respective thickness of the arrows is proportional to the number of events observed.

When disregarding the direction of the interaction as in Figure 9 we see that most of the events between the three actors describe material conflict or verbal cooperation as defined by the quad scale.

A.3 Bazzi et al. (2021)'s structural variables

The dataset constructed by Bazzi et al. (2021) includes a rich set of covariates on the demography and economic activity, the geography and weather, historical variables, political variables, the municipality distances to the Caguan demilitarized zone (DMZ), and

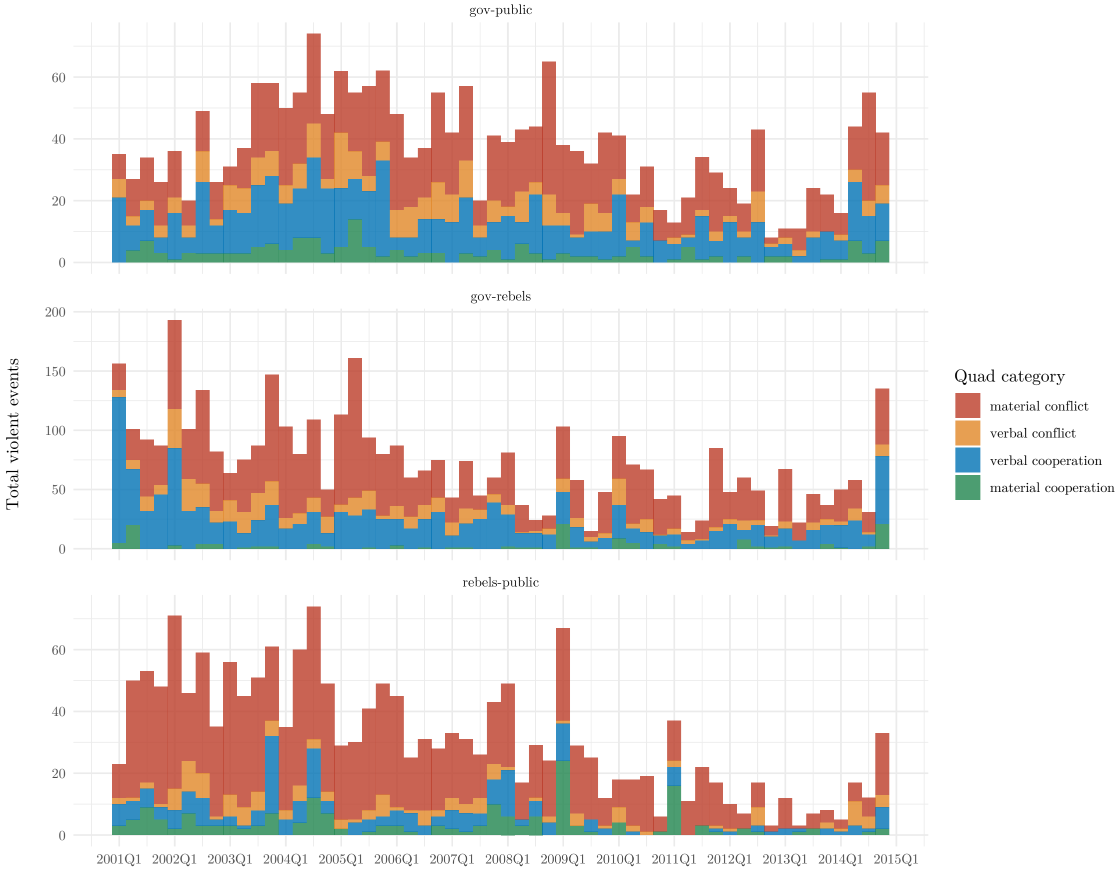


Figure 9: The plots show the count of ICEWS events for all municipalities across the period of the study. The event count is split by actor pair and quad category of the event. On the one hand we see that most of the interactions happened between the government and the rebels while on the other hand we see that events relating to material conflict were the most frequent ones across the actor pairs.

US military aid. A complete overview is available in the online Appendix of Bazzi et al. (2021). The variables considered as structural include variables on the demography and economic activity, the geography and weather, historical variables and the DMZ distances. Since not all of the variables are time invariant but slow moving (population data for instance) instead, we take the mean of the latter between 1991 and 2014 in order to erase any time variation that might capture non-structural features.

The dataset contains data on 1023 distinct municipalities taken from the Dube and Vargas (2013) dataset. While most if the identifiers are the DANE municipality codes, some are

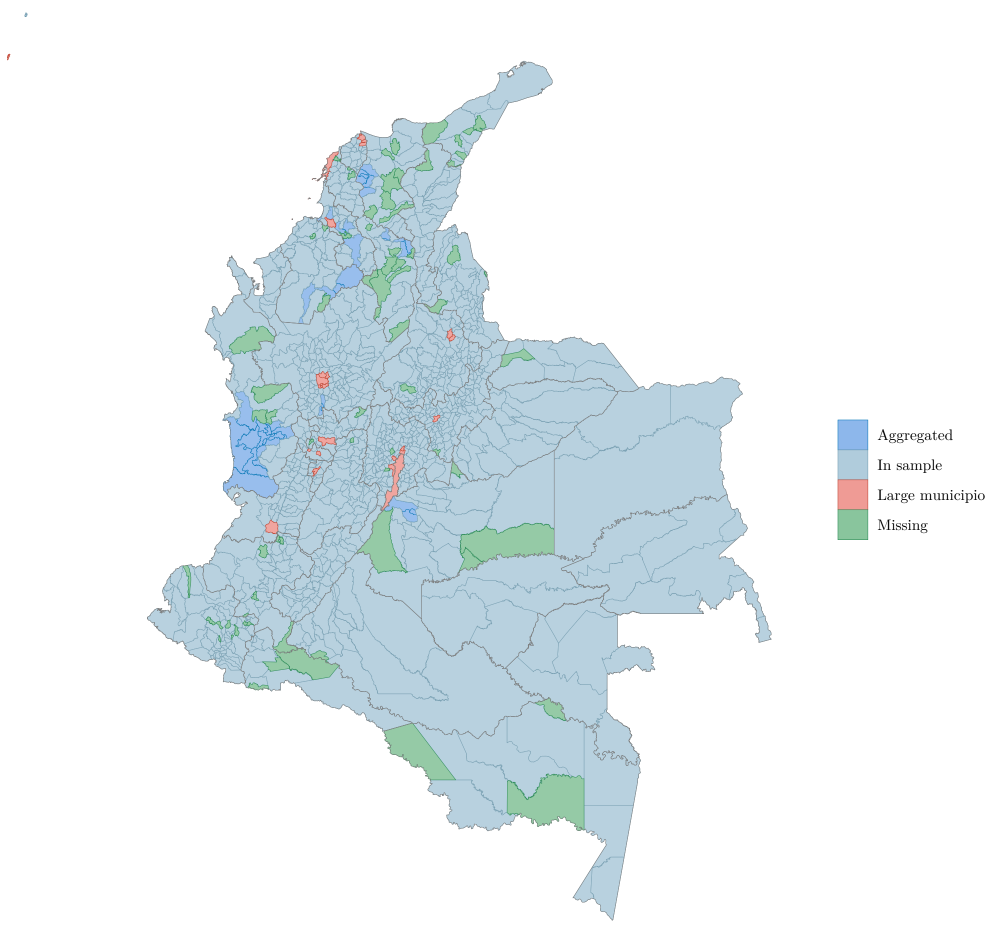


Figure 10: The map indicates the status of the different Colombian municipalities in our study.

aggregates of multiple municipalities that were created after 1991.[29] Other municipalities that were created between 1991 and 2005 were not included in the study. As a result, we end up with a sample of the municipalities shown in Figure 10.

[29] The different components of the aggregated DANE codes are available in the ReadMe file to the replication package of the Dube and Vargas (2013) paper.

B Methodology

B.1 Random Forest

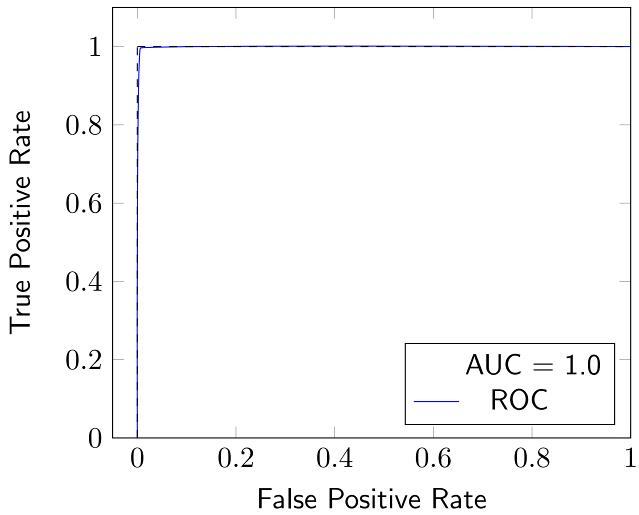
A Random Forest can be considered as an aggregate of a collection of random and independent decision trees. A decision tree splits a sample into subgroups through recursive partitioning and consists of a root, nodes, splits and leafs. From the starting point, the root, the dataset is split into subgroups called nodes. The splits are done through variable cutoffs that minimize an impurity criterion which is most commonly the Shannon entropy or the Gini index, where the former can be interpreted as maximizing the information gain and the latter as maximizing the variance that is explained by the variable used for splitting. The terminal nodes are also referred to as leafs and the prediction for each leaf is the mean outcome for the observations at that leaf. Decision trees are non-parametric and robust to noisy variables and outliers and furthermore allow for conducting dimensionality reduction since only the most relevant variables are considered but tend to overfitting (high variance low bias), however. (Louppe, 2014)

To overcome this overfitting of the individual trees the Random Forest relies on bootstrap aggregating (bagging) and random feature selection. Bagging ensures that every tree is generated from a randomly selected subsample of the data (with replacement) and the random feature selection randomly limits the potential candidates for a split at every node. The number of variables proposed for a split at each node (P) has to be chosen manually. For a classification problem the prediction is done through majority voting of the individual tree predictions.

Random Forests can also be modified to account for imbalanced data, a feature that is especially important for violent outcomes that are typically unbalanced as we usually have more peaceful observations than violent ones. Imbalanced data is insofar problematic as it can lead to the bootstrap sample containing few or none outcomes of the minority class which leads to a poor performance in predicting the minority class and most likely results in a bias. The Random Forest can be modified through either cost sensitive learning, where the weight of each class is changed when calculating the impurity score of a chosen split point or through resampling of the bootstrap sample, where either the minority class can be oversampled or the majority class be undersampled. In our setting we choose to downsample the majority class (peaceful outcome).

B.2 Area Under the Receiver Operating Characteristic Curve

The Receiver Operating Characteristic (ROC) curve plots for a binary classification problem the True Positive Rate (TPR)[30], meaning how many positive outcomes were correctly identified, against the False Positive Rate, meaning how many outcomes were falsely predicted as positive. The Area Under the Curve (AUC) measures, as the name indicates, the area under the ROC and can be considered a measure of the ability of a classifier to distinguish between classes. An AUC of 1, as shown in the graph below, represents a perfect forecast while an AUC of 0.5 represents a random guess. A model yielding an AUC of 0.8 is commonly considered as very good while a model yielding an AUC above 0.9 can be considered excellent.



[30]Also known as Recall or Sensitivity

C Alternative Specifications

In order to control whether the inclusion of the sentiment retrieved from the ICEWS dataset is sensitive to the way how the CAMEO intensity score is defined, we derive a fifth specification that considers instead of the various measures for the public sentiment presented in subsection 4.2 event counts of the quad categories by directional actor relationship. The four quad categories are the following: *material conflict*, *verbal conflict*, *verbal cooperation* and *material cooperation*. As for the sentiment measures, we include the counts for the directional relationships as well as the counts of the 'net relationships' with the respective first three lags. The results are presented in Table 2. As we can see are the results from the fifth specification in comparison quite similar to the third specification which allows us to conclude that $\Delta sent_{(p,q)_{i,\,t-1}}$ is not sensitive to the way the CAMEO intensity scores are defined.

We also include a coding alternative to the dependent variable *violence onset* labeled as *violence onset 2*. For the alternative we consider as violence onset a violent period after eight peaceful periods instead of four. Although we observe a difference in the gaps of the performance across the models, we can derive the same performance ranking as before.

Table 2: Results of the one-quarter ahead out of sample forecast using a Random Forest including the quad-count measure and an alternative definition for the violence onset

	Auto-regressive	Structural	Structural + Public Support	Structural + Violence History	Structural + Quad-count Public Support
	(1)	(2)	**(3)**	(4)	(5)
Any violent event	0.855	0.919	0.939	0.939	0.936
$\geq$ 2 violent events	0.903	0.935	0.952	0.953	0.952
Violence onset	-	0.857	0.875	0.827	0.839
Violence onset 2	-	0.845	0.850	0.806	0.853

Notes: The table reports the same results as Table 1 but adds the fifth specification, where $\Delta sent_{(p,q)_{i,\,t-1}}$ is not defined through the CAMEO intensity scores of the ICEWS events but the quad count related to the events instead. It also evaluates the predictive performance of the different models on an alternatively coded specification of the *violence onset* variable.