STATISTICS VS MACHINE LEARNING: AN APPLICATION ON TIME-TO-EVENT DATA IN TERRORIST KIDNAPPING EVENTS

Luciano Nieddu, Cecilia Vitiello

Abstract. Hostage-taking durations in terrorist attacks using the Global Terrorism Database have been considered in this paper using data from 1970 to 2021. Employing conditional inference trees and Cox proportional hazards models, we managed to determine factors influencing hostage release time. Attack type and ransom demands significantly prolong incident duration, with effects varying over time. Regional variations have also been detected, with Middle East & North Africa and Southeast Asia showing longer median durations. The complementary insights from machine learning providing a clustering of individuals based on survival time and statistical methodologies which provide a clustering of the effects of the covariates on the instantaneous risk of surviving yield a robust framework for understanding complex events such as kidnapping.

1. Introduction

The Global Terrorism Database (GTD) (LaFree and Dugan, 2007; START, 2022) is one of the most extensive and systematically compiled longitudinal dataset on terrorist events. GTD covers events from 1970 to 2021, except for 1993 due to data loss, and includes over 214000 recorded attacks. The database is constructed from open-source materials such as media reports and official documents and provides information on various dimensions of each incident, covering event-level information such as the date, location, weapons used, target type, and number of fatalities and injuries. The GTD has been widely employed in empirical research on terrorism, including statistical modeling of temporal and spatial trends (Dugan et al, 2005; Phillips, 2011), and assessments of counter-terrorism policy effectiveness (Enders and Sandler, 2012). Several studies have used GTD data to investigate tactical selection and target profiles (Deloughery et al. 2012), the impact of terrorist negotiations and concessions (Piazza, 2017), and the effectiveness of rescue operations and government responses (White et al 2014, Wilson, 2000).

The longitudinal nature of GTD makes it particularly well-suited for time series modeling and survival analysis. In this study, we consider the hostage subset of GTD records to quantify risk factors and detect patterns in kidnapping events over time and across geographic regions. Several studies have examined the duration of kidnappings: Brandt and Sandler (2009) analyzed the dynamic properties of hostage-taking series using time series methods, finding that past concessions influence future kidnapping events. Kim and Sandler (2021) applied survival analysis to a dataset of terrorist hostage-taking incidents from 1978 to 2018 identifying key determinants that prolong the duration of hostage situations (i.e. demanding release of imprisoned comrades). Gaibulloev and Sandler (2009) investigates the determinants of logistical and negotiation success in terrorist hostage-taking missions finding that incident durations are positively associated with negotiation success. Phillips (2014) used survival analysis to investigate the factors contributing to the varying longevity of terrorist groups.

This paper uses machine learning and semi parametric survival analysis techniques to assess the instantaneous risk of a hostage release as a function of several structural characteristics of the attack.

2. Data and Methods

Out of the total number of events only those classified as "hostage kidnapping" (i.e. held against their will) or "kidnapped" (i.e. held against their will and taken to another location) have been selected resulting in a dataset of 17161 cases. Duration of the kidnapping has been measured in hours since the incident. Cases where the duration was not available have been removed from the dataset resulting in a dataset of 5534 cases. Events have been right-truncated to the 95th percentile of all the times resulting in a dataset of 5258 observations. Events happening at time zero (240) or with a negative duration (3) were removed resulting in a dataset of 5015 records.

The event has been defined as the release of the hostage within the observational period. Hostages not released or killed have been considered as censored. In each attack the number of persons taken hostages and the number of persons released have been used as weights for the analysis. Those cases where the weight was not available were removed (162) resulting in a final dataset of 4853 cases.

The survival function $S(t)$, modelling the time-to-event of the hostage, has been estimated via the product limit Kaplan Maier method (Kaplan and Meier, 1958) and the semi-parametric Cox proportional hazards model (Cox, 1972). The results of these statistical techniques have been compared with the machine learning approach based on survival trees (Hothorn et al. 2006; Hothorn and Zeileis 2015). The variables that have been considered int the models are:

- Attack Type (attacktype): primary tactical method of the incident (e.g., armed assault, bombing, hostage-taking).
- Region (region): geographical region where the incident occurred.

- Ransom (ransom): binary variable indicating whether a monetary ransom was demanded during the incident.
- Suicide (suicide): binary variable indicating if the attack was a suicide operation.
- Success (success): binary indicator of the attack's outcome, denoting whether the violent act was executed as intended (e.g., a bombing is successful if the device detonates; an assassination is successful only if the specific target is killed).

The proportionality assumption for the hazard related to the Cox model has been assessed via the Shoenfeld residuals. For those variables whose effect does not meet the proportionality assumption, time-dependent effects were included in the model by introducing the interactive effect between the variables themselves and the log(t).

2.1. Kaplan Meier Product limit estimator

The Kaplan-Meier (KM) product limit estimator is a nonparametric method for estimating the survival function from time-to-event data in the presence of censored observations. Developed by Kaplan and Meier (1958), the estimator computes the probability of survival beyond specific time points by multiplying conditional survival probabilities observed at each distinct event time. It does not require any assumptions about the underlying hazard function or the effect of covariates, making it especially useful for descriptive analysis. Censored observations are accounted for by reducing the number at risk when the censoring occurs without contributing to the event count.

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right)$$

Where:

- t_i represents the ordered distinct event times ($t_1 < t_2 < \cdots < t_k$).
- d_i is the number of events (deaths or failures) observed at time t_i.
- n_i is the number of individuals at risk (alive and uncensored) just prior to time t_i.

2.2. Cox Model

The Cox proportional hazards model is a semi-parametric method used to analyze time-to-event data, originally introduced by Cox (1972). It estimates the hazard function, i.e., the instantaneous risk of the event, conditional on a set of covariates. The model assumes that covariates have a multiplicative effect on the baseline hazard and that these effects are constant over time (proportional hazards assumption).

$$h(t) = h_0(t) \cdot e^{X\beta}$$

Estimation is performed via partial likelihood, which allows the baseline hazard to remain unspecified while estimating the regression coefficients. The estimates of the β coefficients are the conditional effect of each single covariate on the hazard on the log scale.

2.3. Survival Trees

Survival trees used in this paper are based on conditional inference trees. The method is based on the conditional inference framework developed by Hothorn et al (2006), which uses statistical hypothesis testing to grow a binary tree. The global null hypothesis of independence between each covariate and the survival response is evaluated using the log-rank test. If a significant association is detected, the variable with the strongest evidence against the null hypothesis is selected, and the best binary split is identified to maximize the separation in survival experiences between resulting groups. The recursive process continues until no covariate exhibits a significant relationship with the survival outcome. At each terminal node, the model yields an estimated survival function (Kaplan-Meier) for the statistical units assigned to that terminal node. Conditional inference trees for survival data offer a robust, nonparametric method to identify covariates influencing time-to-event outcomes, especially in the presence of complex interactions and non-linear effects.

3. Results

From the KM estimator, the median event durations provide a robust measure of central tendency (see Table 1), less susceptible to the influence of extreme values. Central Asia, Middle East & North Africa (MENA), and Sub-Saharan Africa all report a median duration of 120 hours (Table 2). This group primarily consists of regions that have historically experienced complex geopolitical conditions, internal conflicts, or varying degrees of institutional instability. Numerous regions, such as Australasia & Oceania, Central America & Caribbean, and East Asia, present a median duration of 24 hours, suggesting a prevalence of relatively short events.

Table 1 – *KM statistics for survival time.*

N	n.start	events	rmean	se(rmean)	median	LCL	UCL
38476	38476	32685	335.46	4.58	72	72	72

The median survival time for the entire dataset is 72 hours from the kidnapping (see Table 1). In Figure 1 the survival tree output has been displayed. The four terminal nodes identify 4 clusters of regions with homogenous behavior according to the time-to-event data. Node 6 (MENA and Southeast Asia) is the one presenting the cases more at risk of presenting a long duration of the kidnapping compared to the other 3 clusters. The cluster showing the fastest resolution time is represented by node 3 (Australasia and Oceania, Central America and Caribbean, Central Asia, East Asia, Western Europe).

Table 2 – *KM statistics using region as covariate.*

Region	N	n.start	events	rmean	se(rmean)	median	0.95LCL	0.95UCL
A/O	47	47	44	73.02	15.20	24	16	24
CA/Car	488	488	459	98.45	14.39	24	24	24
C.Asia	243	243	218	197.43	32.91	120	72	120
E.Asia	275	275	274	24.18	0.27	24	24	24
E.E.	2847	2847	2330	152.76	10.66	72	72	72
MENA	8151	8151	6002	698.89	16.92	96	96	96
N.Am.	647	647	376	93.44	19.54	30	30	48
S.Am.	3082	3082	2816	315.49	14.84	48	48	48
S.Asia	9346	9346	7872	215.50	6.69	48	48	48
SE.Asia	2866	2866	2673	506.01	19.22	48	48	72
SSA	8745	8745	7995	270.16	6.15	120	120	120
W.E.	1739	1739	1626	179.64	9.65	37	37	37

Note: Australasia & Oceania (A/O), Central America & Caribbean (CA/Car), Central Asia (C.Asia), East Asia (E.Asia), Eastern Europe (E.E.), Middle East & North Africa (MENA), North America (N.Am.), South America (S.Am.), South Asia (S.Asia), Southeast Asia (SE.Asia), Sub-Saharan Africa (SSA), Western Europe (W.E.)

Figure 1 – *output of survival tree using region as covariate.*

```
[1] root
|   [2] region in A/O, CA/Car, C.Asia, E.Asia, E.E., N.Am., SSA, W.E.
|   |   [3] region in A/O, CA/Car, C.Asia, E.Asia, W.E.: 72.000 (n = 279)
|   |   [4] region in E.E., N.Am., SSA: 120.000 (n = 987)
|   [5] region in MENA, S.Am., S.Asia, SE.Asia
|   |   [6] region in MENA, SE.Asia: 288.000 (n = 1434)
|   |   [7] region in S.Am., S.Asia: 192.000 (n = 2153)

Number of inner nodes:    3
Number of terminal nodes: 4
```

Note: Australasia & Oceania (A/O), Central America & Caribbean (CA/Car), Central Asia (C.Asia), East Asia (E.Asia), Eastern Europe (E.E.), Middle East & North Africa (MENA), North America (N.Am.), South America (S.Am.), South Asia (S.Asia), Southeast Asia (SE.Asia), Sub-Saharan Africa (SSA), Western Europe (W.E.)

The Cox proportional hazards model provides a quantification of the effect on the log-hazard of each value of the covariate while the survival tree provides a clustering of the regions based on the survival probability. Estimates for the Cox model using region as covariate have been reported in Table 3.

Comparing the groups of regions obtained by the survival tree approach (Figure 1 and 2) with the effect on the log hazard (Table 3) we observe:

- Node [3]: fastest release group (Median 72 hours): Australasia & Oceania (Reference), Central America & Caribbean, Central Asia, East Asia, Western Europe. *The Cox model shows that regions in this group show non significant effect on the log(HR)*
- Node [4]: moderately fast release Group (Median 120 hours): Eastern Europe, North America, Sub-Saharan Africa. *The Cox model's results for these regions present moderately negative log(HR) values that are either statistically significant or borderline.*
- Node [7]: fairly slow-release group (Median 192 hours): South America, South Asia. *The Cox model presents significantly negative log(HR) values for these regions although not the highest observed effects.*
- Node [6]: slow-release group (Median 288 hours): MENA, Southeast Asia. *The regions in this node show the highest and statistically significant negative log(HR) values in the Cox model.*

Table 3 – *Cox proportional hazard model estimates.*

Characteristic	*log(HR)*	*95% CI*	*p-value*
region			
Australasia & Oceania	—	—	
Central America & Caribbean	***-0.57***	***-1.3, 0.20***	***0.150***
Central Asia	***-0.53***	***-1.4, 0.37***	***0.300***
East Asia	***0.58***	***-0.51, 1.7***	***0.300***
Western Europe	***-0.71***	***-1.5, 0.05***	***0.069***
MENA	*-1.3*	*-2.0, -0.55*	*<0.001*
Southeast Asia	*-1.3*	*-2.1, -0.56*	*<0.001*
South America	*-1.2*	*-1.9, -0.43*	*0.002*
South Asia	*-1.1*	*-1.8, -0.36*	*0.004*
Eastern Europe	*-0.87*	*-1.6, -0.10*	*0.026*
North America	*-0.80*	*-1.6, 0.00*	*0.051*
Sub-Saharan Africa	*-0.91*	*-1.7, -0.17*	*0.017*

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Using the clustering of regions suggested by the survival tree we observe that regions belonging to Node 3 are those regions in the Table 3 that show not significant or marginally significant effects on the risk of the event. Regions in node 4 show significant and average negative effect on the risk to experience the event. Regions from Node 6 show very strong negative significant effect on the risk of experiencing the event and regions in Node 7 show strong middle high negative significant effect

3.1. Extended Model

The goal of this section is to compare the performances of the machine learning approach and the semi-parametric approach on a model where the hazard function depends on more than one variable. Namely the following model has been considered:

$$S(t) = f(Region, Attacktype, Ransom, Suicide, Success)$$

where:

- *Region* is the geographical region where the event took place
- *Attacktype* is a variable detailing the general method of attack. It takes value as “Assassination”, “Hijacking”, “Kidnapping”, “Barricade Incident”, “Bombing/Explosion”, “Armed Assault”, “Unarmed Assault”, “Facility/Infrastructure Attack”, “Unknown”;
- *Ransom* is a binary variable which is 1 iff there is evidence that a ransom was requested
- *Suicide* is a binary variable which is 1 iff there is evidence that the terrorists were planning to commit suicide
- *Success* is a binary variable which is 1 if the attack was successful

The survival tree yields a final partition of the statistical units into nine terminal nodes (Figure 2). The first splitting has been performed using type of attack. while the following split involves different choices of variables implying a different role of the variables within subsets of the upper level variables. This suggests a possible presence of interaction and nonlinear effect of covariates on the outcome.

The survival tree suggests once again a partition of the regions of interest into two major subsets, namely:

- Australasia & Oceania, Central America & Caribbean, Central Asia, East Asia, Eastern Europe, North America, Sub-Saharan Africa, Western Europe
- MENA, South Asia, Southeast Asia

In Figure 2 the median survival times for each terminal node have been reported together with the size of the terminal node; “Inf” values for the median indicate that the cumulative incidence in that cohort of cases remained below 50% making impossible the estimation of the median time. This identifies a small subgroup with a relatively low hazard rate, and the follow-up time is insufficient to observe the point at which half of the cohort has experienced the event.

Figure 2 – *Output of survival trees using all covariates.*

```
[1] root
|   [2] attacktype in Assau, Assas, Kidnap, Other
|   |   [3] region in A/O, CA/Car, C.Asia, E.Asia, E.E., N.Am., SSA, W.E.
|   |   |   [4] attacktype in Assau, Kidnap: 120.000 (n = 1096)
|   |   |   [5] attacktype in Assas: 576.000 (n = 45)
|   |   [6] region in MENA, S.Am., S.Asia, SE.Asia
|   |   |   [7] region in MENA, SE.Asia
|   |   |   |   [8] attacktype in Assau, Assas
|   |   |   |   |   [9] ransom <= 0
|   |   |   |   |   |   [10] region in MENA: Inf (n = 32)
|   |   |   |   |   |   [11] region in SE.Asia: 432.000 (n = 20)
|   |   |   |   |   [12] ransom > 0: 4320.000 (n = 10)
|   |   |   |   [13] attacktype in Kidnap, Other: 288.000 (n = 1302)
|   |   |   [14] region in S.Am., S.Asia: 192.000 (n = 2076)
|   [15] attacktype in Expl, Inf.Att, Hijack, Barr
|   |   [16] attacktype in Expl, Hijack: 48.000 (n = 124)
|   |   [17] attacktype in Inf.Att, Barr: 10.000 (n = 148)

Number of inner nodes:    8
Number of terminal nodes: 9
```

Note: when a median survival time is reported as (Inf) (infinity) for a terminal node, it indicates that the median survival time for subjects in that specific subgroup was not reached during the study's follow-up period.
Regions: A/O: Australasia & Oceania, CA/Car: Central America & Caribbean, C.Asia: Central Asia, E.Asia: East Asia, E.E.: Eastern Europe, MENA: Middle East & North Africa, N.Am.: North America, S.Am.: South America, S.Asia: South Asia, SE.Asia: Southeast Asia, SSA: Sub-Saharan Africa, W.E.: Western Europe.
Attack Types: Assau: Armed Assault, Assas: Assassination, Kidnap: Hostage Taking (Kidnapping), Other: Unknown, Expl: Bombing/Explosion, Inf.Att: Facility/Infrastructure Attack, Hijack: Hijacking, Barr: Hostage Taking (Barricade Incident)

The large medians in some subgroups (4320 hours in Node 12 for instance) can be attributed to the recursive partitioning of the covariate space: the tree algorithm has isolated a very specific and small subset of the data (n=10 for Node 12) defined by the pathway leading to that node. This partition identifies a cohort with a relatively low hazard rate.

The Cox model has been fitted starting from the null model with just the intercept and using a stepwise forward selection strategy to improve the model fitting using penalized likelihood (AIC).

$$h(t|X) = h_0(t)e^{(\beta_1 \cdot \text{attacktype} + \beta_2 \cdot ransom + \beta_3 \cdot \text{region} + \beta_4 \cdot success + \beta_5 \cdot suicide)}$$

The covariates in the final model are shown in Table 4 according to the order of inclusion.

Table 4 – *Cox model forward selection.*

Step	d.o.f.	Deviance	Resid. Df	Resid. Dev	AIC
	--	--	3158	46390.86	46390.86
+Attacktype	-7	326.45654	3151	46064.41	46078.41
+Region	-11	100.99569	3140	45963.41	45999.41
+Ransom	-1	30.68959	3139	45932.72	45970.72

The final Cox proportional hazards model selected by the stepwise procedure is:

$$h(t|X) = h_0(t)e^{\beta_1 \cdot \text{attacktype} + \beta_2 \cdot \text{region} + \beta_3 \cdot \text{ransom}}$$

The proportionality assumption does not hold for *Ransom* and *Attacktype*, therefore Attacktype has been recoded as $Attackidnap$ which is equal to one *iff* Attacktype="Hostage Taking (Kidnapping)", i.e. those attacks that originally are conceived with the idea of taking hostages. A time dependent effect has been used for both Attackidnap and Ransom allowing therefore the effects of these two variables to change over time. The new model is therefore:

$$h(t|X) = h_0(t)e^{(\beta_1 \cdot \text{attackidnap} + \beta_2 \cdot \log(t) \cdot \text{attackidnap} + \beta_3 \cdot \text{region} + \beta_4 \cdot \text{ran} + \beta_5 \cdot \log(t) \cdot \text{ran})}$$

All other things being equal, Attackkidnap (Table 5) shows a negative significant influence on the instantaneous hazard of incident resolution (log(HR) = -1.7; p<0.001), i.e. a significantly protracted duration for these specific events. The time-dependent effect for Attackkidnap is also highly significant and positive (log(HR) = 0.36; p<0.001) therefore while pure hostage-taking attacks generally prolong incident duration, this effect is dampened as time progresses, all other things being equal.

The presence of a ransom note or a ransom demand shows a similar behavior: highly significant negative association with the hazard of resolution (log(HR) = -0.75; p<0.001). This finding is coherent with the expectation that ransom negotiations introduce initial delays. As for Attackidnap the time-dependent effect for ransom is statistically significant and positive (log(HR) = 0.11; p<0.001).

Considering the regional effect on the instantaneous hazard of incident resolution the model suggests different patterns across different geographical regions, (reference category: Australasia & Oceania). Namely: a significant negative effect on the hazard for a group of regions where the kidnapping attacks tend to be of longer duration. This group includes:

- MENA (log(HR) = -1.1, p-value = 0.003)
- Southeast Asia (log(HR) = -1.1, p-value = 0.003)
- South America (log(HR) = -1.0, p-value = 0.008)

- South Asia (log(HR) = -0.97, p-value = 0.010)

In these regions, the instantaneous hazard of resolution is reduced compared to Australasia & Oceania suggesting a significantly longer time-to-event. Specifically, the hazard is about 67% lower in the MENA and Southeast Asia regions, 63% lower in South America, and 62% lower in South Asia compared to Australasia & Oceania. This pattern suggests potentially shared regional dynamics, varying levels of state response capacity, or differing cultural/political approaches to hostage situations that can contribute to more prolonged kidnapping time.

Table 5 – *Cox proportional hazard model estimates (forward selection).*

Characteristic	*log(HR)*	*95% CI*	*p-value*
attackidnap	*-1.7*	*-1.9, -1.4*	*<0.001*
tt(attackidnap)	*0.36*	*0.29, 0.43*	*<0.001*
factor(region)			
Australasia & Oceania	—	—	
Central America & Caribbean	*-0.38*	*-1.2, 0.39*	*0.300*
Central Asia	*-0.48*	*-1.4, 0.42*	*0.300*
East Asia	*0.44*	*-0.66, 1.5*	*0.400*
North America	*-0.70*	*-1.5, 0.11*	*0.089*
Western Europe	*-0.53*	*-1.3, 0.23*	*0.200*
Eastern Europe	*-0.78*	*-1.6, -0.01*	*0.046*
Sub-Saharan Africa	*-0.76*	*-1.5, -0.01*	*0.046*
MENA	*-1.1*	*-1.9, -0.40*	*0.003*
South America	*-1.0*	*-1.8, -0.26*	*0.008*
South Asia	*-0.97*	*-1.7, -0.23*	*0.010*
Southeast Asia	*-1.1*	*-1.9, -0.38*	*0.003*
ran	*-0.75*	*-1.0, -0.49*	*<0.001*
tt(ran)	*0.11*	*0.06, 0.16*	*<0.001*

Another group are those regions that show a marginally significant effect on the log hazard

- Eastern Europe (log(HR) = -0.78, p-value = 0.046)
- Sub-Saharan Africa (log(HR) = -0.76, p-value = 0.046)

These regions show a reduction in hazard of approximately 53% to 54% relative to Australasia & Oceania. These effects, while significant, are of a smaller magnitude than those in the previous group.

The last group is those of regions that do not show any statistically significant effect compared to the reference category (Australasia & Oceania):

- Central America & Caribbean (p-value = 0.300),
- Central Asia (p-value = 0.300),
- East Asia (p-value = 0.400),

- Western Europe (p-value = 0.200),
- North America (p-value = 0.089)

Kidnapping incident durations in these regions are not statistically distinguishable from those in Australasia & Oceania all other things being equal.

4. Conclusions

This study applied both conditional inference trees and the Cox proportional hazards model to analyze survival times in terrorist events involving hostages. Conditional inference trees offered a non-parametric framework to simultaneously perform feature selection, identify clusters, and handle both qualitative and quantitative variables without requiring distributional assumptions. They overcome the difficulty of Kapan-Meier product limit estimator to efficiently handle quantitative covariates or qualitative covariates with a large number of categories. The Cox model, as a semi-parametric method, provided an inferential framework that allowed for quantitative estimates of hazard ratios, testing of covariate effects and varying covariate effects over time. It supported feature selection and clustering of events based on estimated coefficients and their statistical significance. The Cox model findings quantified the impact of various factors on event duration, highlighting that ransom demands are significantly associated with longer survival times.

The results obtained from the tree-based and Cox model approaches were coherent and complementary. While the Cox model offered effect estimates and statistical inference, machine learning trees suggested complex interactions and clusters in terrorist behavior. These methods can be used in cascade (either tree and KM or tree and cox model), where exploratory insights from the tree model can inform the specification of the Cox model or suggest categorization of quantitative variables for the Kaplan-Meier. The congruence between the Cox model's parameter estimates (magnitude and significance of log(HR)) and the regional partitions obtained by the survival tree provides evidence that geographical region is a fundamental determinant of hostage release duration.

References

BRANDT P.T., SANDLER T. 2009. Hostage taking: Understanding terrorism event dynamics, *Journal of Policy Modeling*, Vol. 31, No. 5, pp. 758–778.

COX D.R. 1972. Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 34, No. 2, pp. 187–220.

DELOUGHERY K., KING R.D., ASAL V. 2012. Close cousins or distant relatives? The relationship between terrorism and hate crime, *Crime & Delinquency*, Vol. 58, No. 5, pp. 663–688.

DUGAN L., LAFREE G., PIQUERO A.R. 2005. Testing a rational choice model of airline hijackings, *Criminology*, Vol. 43, No. 4, pp. 1031–1065.

ENDERS W., SANDLER T. 2012. The Political Economy of Terrorism. Cambridge: Cambridge University Press.

GAIBULLOEV K., SANDLER T. 2009. Hostage taking: Determinants of terrorist logistical and negotiation success, *Journal of Conflict Resolution*, Vol. 46, No. 6, pp. 882–901.

HOTHORN T., HORNIK K., ZEILEIS A. 2006. Unbiased recursive partitioning: A conditional inference framework, *Journal of Computational and Graphical Statistics*, Vol. 15, No. 3, pp. 651–674.

HOTHORN T., ZEILEIS A. 2015. partykit: A Modular Toolkit for Recursive Partytioning in R, *Journal of Machine Learning Research*, Vol. 16, pp. 3905–3909.

KAPLAN E.L., MEIER P. 1958. Nonparametric estimation from incomplete observations, JASA, Vol. 53, No. 282, pp. 457–481.

KIM W., SANDLER T. 2021. Non-UN Peacekeeping Effectiveness: Further Analysis, *Defense and Peace Economics*, Vol. 33, No. 5, pp. 511–533.

LAFREE G., DUGAN L. 2007. Introducing the Global Terrorism Database, *Terrorism and Political Violence*, Vol. 19, No. 2, pp. 181–204.

PHILLIPS B.J. 2014. Terrorist group cooperation and longevity, *International Studies Quarterly*, Vol. 58, No. 2, pp. 336–347.

PHILLIPS P.J. 2011. The Diseconomies of Terrorism, *World Economics*, Vol. 12, No. 4, pp. 171–192.

PIAZZA J.A. 2017. Repression and terrorism: A cross-national empirical analysis of types of repression and domestic terrorism, *Terrorism and Political Violence*, Vol. 29, No. 1, pp. 102–118.

START - National Consortium for the Study of Terrorism and Responses to Terrorism. 2022. Global Terrorism Database (GTD). Retrieved from http://www.start.umd.edu/gtd

WHITE G., MAZEROLLE L., PORTER M., CHALK P. 2014. Modelling the effectiveness of counter-terrorism interventions, *Trends & Issues in Crime and Criminal Justice*, No. 475. Canberra: Australian Institute of Criminology.

WILSON M.A. 2000. Toward a model of terrorist behavior in hostage-taking incidents, *Journal of Conflict Resolution*, Vol. 44, No. 4, pp. 403-424.

Luciano NIEDDU, UNINT University, l.nieddu@unint.eu
Cecilia VITIELLO, Sapienza University of Rome, cecilia.vitiello@uniroma1.it