

SAGE Research Methods Cases Politics & International Relations Submission for Consideration

Case Title

A Case Study of Spatial Analysis: Approaching a Research Question with Spatial Data

Author Name(s)

Andre Python¹ & Jürgen Brandsch²

Author Affiliation & Country of Affiliation

¹University of Oxford, Nuffield Department of Medicine, Big Data Institute, Oxford, UK

²University of Mannheim, Department of Economics, Mannheim, Germany

Lead Author Email Address

Email: andre.python@bdi.ox.ac.uk

Discipline: D8 [*please do not alter*]

Sub-discipline within Politics & International Relations

International Relations [SD-POLIR-5]

Academic Level of intended readership

Postgraduate

Contributor Biographies

Andre Python is a Post-Doctoral Research Scientist in Geospatial Epidemiology in the Malaria Atlas Project at the Big Data Institute, University of Oxford. His first degree was in Geography (University of Fribourg, Switzerland). In 2017, he completed his PhD in Statistics

(University of St Andrews, UK), applying a Bayesian spatio-temporal geostatistical model to capture fine-scale patterns of non-state terrorism across the world. Andre's current interests are in extending geostatistical models to address issues posed to global health and security.

Jürgen Brandsch is a Post-Doctoral Researcher at the University of Mannheim, where he works in the ERC-funded project "Repression and the Escalation of Violence" (RATE). Jürgen's general research interests focus on conflict dynamics, violence against civilians, terrorism and ethnic conflicts. Before joining RATE Jürgen undertook his doctoral studies at the University of St Andrews and worked for several years for the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), primarily in the field of Peace and Security. Jürgen also holds a Diplom in Economics (University of Heidelberg) and Magister Artium in Political Science (University of Münster).

Published Article

Python, A., Brandsch, J., & Tskhay, A. (2017). Provoking local ethnic violence—A global study on ethnic polarization and terrorist targeting. *Political Geography*, 58, 77-89.

Abstract

Research work that aims at capturing fine-scale patterns of natural or social phenomena require spatially accurate data and the use of suitable methodological tools. The increasing availability of data gathered at finer spatial scales and the advent of geographical information systems (GIS) packages along with spatial statistical modelling tools have made possible the investigation of natural, and more recently, social phenomena at fine spatial scales. In the field of Politics and International Relations, these advances have offered new frameworks to assess research hypotheses and allowed scholars to provide results spatially tailored to the

requirements needed to aid policy makers' decisions on the ground. Yet the GIS packages and statistical models that handle and analyse spatial datasets—sub-national data often include a very large number of observations—have steadily increased in complexity and their usage have posed new challenges to non-specialists. In this work, we identify and describe five main issues that arose during a research process carried out to better understand the link between ethnicity and terrorism at a fine spatial scale. Our case study is framed within a very specific context and offers one among many alternative approaches to tackle the challenges posed by spatial data. Yet we identified and provided guidance and recommendations to address key issues that are likely to be encountered in a wide range of research work using spatial data.

Learning Outcomes

“By the end of this case, students should be able to:

- Identify four key elements related to the use of spatial data:
 1. modelling approach (point, lattice, geostatistics);
 2. spatial scale;
 3. spatial aggregation;
 4. probability distribution of the dependent variable.
 - Understand the connections among these elements
 - Adapt each element to answer the research question
-

A Case Study of Spatial Analysis: Approaching a Research Question with Spatial Data

Introduction

Project Overview

Terrorism has become a major security threat over the last 50 years. The National Consortium for the Study of Terrorism and Responses to Terrorism (2017) reported that more than 170,000 terrorist attacks have been carried out by non-state actors across the world from 1970 to 2016. However, terrorism is not distributed equally among countries regions or cities. While it is currently a relatively rare phenomenon in Western countries other states like Afghanistan and Iraq experienced massive amounts of violence over the last decade. Even in these countries some locations are more often targeted than others and researchers have recently begun to focus on understanding terrorism within states.

In general, scholars acknowledge that the causes of terrorist attacks are highly complex and operate at various scales, encompassing the individual motives of terrorists, organizational abilities to structural factors and spill-over effects (English, 2010; Hoffman, 2006; Richardson, 2007). From a Spatial Statistics point of view, this multi-scale nature of terrorism implies that the association between *covariates*, also-called independent or explanatory variables (e.g. population density, political system) and the *dependent variable* (e.g. number of terrorist attacks) may vary across spatial scales. In (2016) Enders and his colleagues brought evidence that terrorist attacks are more frequent in countries with mid-range per capita gross domestic product (GDP), and less frequent with low and high per capita GDP countries, thereby suggesting a nonlinear (an “inverse-U-shape”) relationship. However, Python and his colleagues (2018) show that a linear relationship cannot be excluded at subnational level. It thus depends on the research question, whether one should study

terrorism as a country-based or as a more localised phenomenon. In practice, this means choosing a suitable *spatial unit of observation* (i.e. the spatial level at which data are aggregated), which allows researchers to:

- capture patterns of a phenomenon at a desired spatial scale;
- assess theories in line with the spatial scale of their predictions;
- provide outputs tailored at a suitable spatial scale to be used by policy-makers.

Currently, the most common way to investigate terrorism is to use the country as a spatial unit of observation. For instance, Krueger and Laitin (2008) use country as spatial unit of observation to study the effects of economic and political drivers on terrorism. This type of research makes sense when one asks whether there are country-wide structural drivers that lead people to join a terrorist organization or that lead terrorists to target specific countries. However, all too often the aggregation on the national level obscures pivotal processes that occur at finer scales. For instance, in Brazil, economic development may be high when averaged at a national level but can vary substantially within the country.

The guidance and recommendations described in this present work draw on the research process we have experienced to capture the link between ethnicity and terrorism at a sub-national spatial level of analysis using geographic information systems (GIS) and statistical modelling applied to spatial data. The resulting paper was published in 2017 in *Political Geography*.

Why Going Local?

Terrorism research points out that one of the main strategies of terrorists is to provoke the opposing government. However, we realized that numerous case studies indicated that terrorists may also seek to provoke the local population instead of the government. These case studies seemed to indicate that terrorists tend to attacks areas where it is easy to spark an ethnic or religious conflict that would then help the terrorists to recruit. As a consequence, we

thought of a way to test this hypothesis given the availability of spatial data on terrorism and ethnic groups. Assuming that terrorists employ provocation strategies aiming at deliberately targeting rival ethnic groups to ignite ethnic conflict, we expect terrorist groups to favour ethnically highly polarized areas to maximise their impact.

- *Ethnic polarization* describes one among various facets of ethnic diversity. It is often computed based on the proportion of the area shared by the corresponding ethnic group(s). According to the pioneer work of Reynal-Querol (1998, 2001), peak values of ethnic polarization occur when two ethnic groups of equal size share a given territory.

The reason for this lies in the fact that terrorists need both, an opposition group and their own constituency in the localities of the attack. Only in these cases terrorism can provoke the violent reaction that terrorists are aiming for. Unfortunately, empirical research investigating the role of ethnicity on terrorism has been traditionally carried out at *country level*. Findings at country-level show that countries encompassing a high number of different ethnic groups (Basuchoudhary & Shughart, 2010; Piazza, 2008b) or a high proportion of excluded ethnic groups (Arva & Piazza, 2016; Choi & Piazza, 2016) appear more prone to terrorism.

However, analyses based on data aggregated at country-level are unable to reveal *within-country* variability of ethnic polarization and terrorism and there is surely not a single country where either ethnic settlements or terrorist events are *spatially distributed* in a uniform manner. Thus, studies aggregating at the national level make unwarranted assumptions about the spatial distribution of the investigated phenomena. In addition, it is likely that terrorist events are not spatially independent; rather, they tend to influence each other at close spatial proximity. A spatially disaggregated study on terrorism carried out by Nemeth and his colleagues in 2014 showed that terrorism is often locally *clustered* in space. Thus, the observed within-country variability of terrorism suggests the presence of key factors of

terrorism that vary within-country. This further highlights the need for a subnational investigation of the potential causes of terrorism and requires gathering, aggregating, and analysing data at a local spatial level.

Hence, for theoretical and empirical reasons we decided to study the causes of terrorism locally. Our study is inherently confined to a very specific context and offers one among many valid approaches to tackle the challenges of using spatial data. However, researchers using spatial data face comparable difficulties independently of the context of their study. As we will point out over the next pages, there are many crucial choices to make during the research process and it is not always clear that the one taken is inherently the best. Yet, it is important to reflect on these choices and make them comprehensible for other researchers. This enables research to become a truly collaborative effort allowing fellows to replicate, extend and potentially correct one's own work.

Research Design & Method

I: Defining the Spatial and Temporal Scopes

The first step of our research design consists of defining both the *spatial* and *temporal* scopes of the study, which should be in line with the research question(s) conditional on data availability. Since we were interested in studying terrorism worldwide, we chose a dataset on terrorism that has global coverage. We gathered *spatial points* on terrorism—terrorism data with spatial coordinates (longitude, latitude)—from the *Global Terrorism Database* (GTD) (START, 2016b). In addition, we used *spatial polygons* on ethnic groups—ethnic data aggregated within polygons delimiting their areas of presence—from the *Geo-referencing of ethnic groups* (GREG) (Weidmann, Rød, & Cederman, 2010). To make for ease of comparison among terrorist events, we focused on those which occurred after 9/11 (i.e. from 2002) until 2015, which corresponded to the latest available data at the time of writing.

II: Defining a Suitable Level of Analysis

We outlined in the introductory section that terrorism and its potential association with ethnic polarization is inherently *local*, which calls for a *sub-national* spatial level of analysis. The statistical literature commonly suggests three *statistical approaches* on modelling spatial data. Each approach relates to specific statistical methodologies and relies on a set of different underlying assumptions. The assumption may or may not fit with the data, therefore, selecting an appropriate approach on spatial data requires particular attention. Spatial data can be modelled as:

- *points* (figure 1): the data are *spatially disaggregated* within a study area. This means that each point is provided with a spatial location (e.g. longitude and latitude). In this framework, all points in the study area form one observation, also-called a *point pattern*, which represents one realisation of a *point process*.;
- *geostatistics* (figure 2): the values of the dependent variable (also-called *response*) gathered at different locations represent discrete measures of an underlying spatially *continuous* phenomenon. Therefore, the values of the phenomenon at *unobserved* locations are a priori not known. The observed data is then commonly used to make inference at unobserved locations. For example, one may gather temperature data in cities in the neighbourhood of London to infer temperature values in the centre of the city;
- *lattice* (figure 3): the data is *spatially aggregated* within irregular or regular polygons that cover a study area. Often, aggregated data are reported at the centroid of each polygon:
 - *irregular polygons*: for example, one aggregates data within regional administrative units. In this framework, the shape of the administrative units can vary;

- *regular polygons*: for example, one aggregates data within grid-cells, which correspond to squares of identical size. The size of the grid-cells cannot vary.

-----Insert Figures 1, 2 and 3-----

One possible approach to select a suitable approach on the spatial data is to proceed by elimination, i.e. discarding all approaches whose underlying assumptions cannot reasonably hold. In our work, we did not opt for a point process approach since it is primarily used to model the processes that generated the location of the events. In our case, the location of terrorist events is assumed to be known and our main interest is in examining the effect of ethnic polarization on the number of terrorist events not on their locations. Second, the geostatistical approach assumes that there is an underlying (and unobserved) continuous process that has generated the observed terrorist events. This method may be used to predict terrorism at any location within the study area, which was not the main focus of our study. Third, the lattice approach appears particularly suitable with our research question and the spatial characteristics of our data. Our data could be aggregated within irregular or regular polygons. The aggregation procedure reflects the spatial inaccuracies of our data (see Section III). We favoured the *regular polygons* approach based on grid-cells to ease comparison among spatial units given their identical shape. In contrast, *irregular polygons* replicating the sub-national administrative units would make comparison among units difficult given their different size properties (e.g. in Switzerland, the area of Basel-City Canton is 37 km² while the area of Bern Canton is 5,960 km²).

In addition, regarding the time dimension we made the choice for a reduction of complexity. The phenomenon investigated in our case study can be considered *spatial* rather than *spatio-temporal*. This means that the ethnic settlement patterns that we think influence the location of terrorist attacks are time-invariant—although this may not well reflect the reality, the data are time-invariant. Therefore, we aggregated all variables for regular polygons (grid cells)

over the whole time period so that the time dimension was removed for all variables (see further details in “Research Design & Method - Section IV”).

III: Spatial Accuracy

Having chosen a lattice approach using regular polygons, one needs to define the size of the grid-cells to analyse the investigated process at a desired scale, while accounting for the inevitably limited spatial accuracy of both the dependent variable and the covariates. This is a crucial step since the choice of the grid resolution can have a considerable effect on the results—often known as the *modifiable areal unit problem* (MAUP), well described by Holt and his colleagues (1996). When data are available, one often tries to use the finest grid while ensuring that the statistical analysis is computationally feasible. Like in most empirical studies in social science, spatial inaccuracies in the data prevent us from using very fine grids (unless having recourse to e.g. *data augmentation* method, which is not discussed here. For further interest, see e.g. (Tanner & Wong, 1987)).

First, the spatial accuracy of terrorist events varies considerably. Fortunately, the GTD provides one variable (*specificity*) that allowed us to identify and remove spatially less accurate events according to five levels of spatial accuracy (START, 2016a). In this case we opted to select only the most spatially accurate events (*specificity*=1), which includes terrorist attacks that could be attributed at the city village or town level. In these cases, it is very reasonable to assume that the events are close to the given coordinates (inaccuracy of several kilometres).

Second, the data on the settlement of ethnic groups also incorporates spatial inaccuracies. GREG draws the polygons on ethnicity from the Soviet Atlas Narodov Mira, a 1960s document whose spatial accuracy is not known. In addition, it is highly plausible that some of these settlement patterns have changed over time.

Consequently, it would make little sense to define a very fine-grained spatial unit of observation. We chose PRIO-GRID, which consists of regular grid cells of approximately 55 km by 55 km at the Equator and which is frequently used in conflict research. PRIO-GRID covers all terrestrial areas worldwide (Tollefsen, Strand, & Buhaug, 2012) and we believe that it is the finest spatial scale on which one might capture patterns in the relationship between ethnicity and terrorism using GTD and GREG.

V: Nature and Distribution of the Dependent variable for Spatial Data

Within a lattice framework, the dependent variable represents data aggregated within polygons. In our case, terrorism can be aggregated in several ways. First, we could count the number of attacks, second, we could count the number of fatalities or third we could just identify whether a lethal attack took place. In our case, we chose to count the number of attacks on civilians, irrespective of whether they were lethal or not. In our mind, the lethality of attacks may also hinge on a number of factors such as the competence of the terrorist group, which we are not interested in. In our analysis, the “successfulness” of the attack does not play a role. Therefore, our dependent variable *counts* the number of terrorist attacks (against civilians) aggregated within each grid-cell. This means it is a discrete variable since it only takes whole numbers as possible outcomes, and more specifically, it represents a count-variable as it sums the number of attacks.

The fact that the dependent variable counts events is important as this informs us about the potential *probability distribution* of the variable:

- a probability distribution of the data expresses the probability of occurrence of possible outcomes.

As an example of a continuous probability distribution, the Gaussian distribution, famously known as “bell-curve” due to its bell shape, assigns high probability of occurrence for outcomes close to the mean and a gradual reduction of the probability of occurrence as one

moves away from the mean. However, this usually does not apply to discrete data, such as count of terrorist events. Therefore, one needs to define a probability distribution suitable for count data. Furthermore, count data on terrorism at subnational levels often exhibit a distribution with a high frequency of low values (e.g. 0 or 1 attack) and low frequency of high values (worldwide very few locations experience more than 10 attacks). These properties violate the main assumptions required by an estimation based on the ordinary least squares (OLS). Hence, we opted for the use of the *negative binomial distribution*, which accounts for the discrete nature of the data and the non-normal distribution.

Finally, the fact that we are dealing with spatial data necessitates us to take into account Tobler's *first law of Geography*: "everything is related to everything else, but near things are more related than distant things" (1970). In the spatial statistics jargon, Tobler's law mainly refers to the concept of *spatial autocorrelation*. It occurs when the dependent variable is correlated in space with itself, which violates the assumption of independent events. Therefore, analysts using spatial data need to identify the presence of spatial autocorrelation in the data and to remedy to it. Within a spatial regression context, we used a spatial lag model, which includes means of the dependent variable in neighbouring areas as an extra explanatory variable. In a constantly growing literature, one may find numerous equally valid approaches, including spatial error or geographically weighted regression models to name but two. Choosing an appropriate spatial model is not straightforward and providing a usable guideline would go beyond the scope of this present document.

IV: Robustness Test and Model Selection for Spatial Data

Robustness tests are generally used to assess how the results of a statistical model are affected by changes in its specification. In other words, robustness tests compare outputs obtained from an original model with those from alternative models. One may for example compare the estimated parameter values using an updated version of the dataset, a different

estimation method, or after making changes in the set of covariates. This is very useful to detect whether some relationships between an independent variable and the dependent variable is an artefact to a particular model specification. In our case study, we compared the result of our models using several sets of independent variables, which includes, in particular, ethnic diversity computed through different approaches. Also, we analysed potential changes using different probability distributions associated with the discrete nature of our independent variable. In addition, we compared the models in their ability to predict the value of the dependent variable. Predicted observations can be:

- part of the data sample (*in-sample* predictions) or
- not part of the data sample (*out-of-sample* predictions)

In a spatial modelling framework, changes in the spatial unit of observation may be used as a robustness test or used in the purpose of comparing the predictive performance of models at different spatial scales. In the case of regular lattice data, one may use alternative spatial resolutions of the grid. Choice on the spatial resolution needs to be theoretically justified—the scale on which the phenomenon is observed should match the theoretical framework of the study—and be *sufficiently similar* to the resolution used in the original model to make the comparison meaningful. As a rule of thumb, one may compare the results using four alternative grids with resolutions reduced by 10% and 5% and increased by 5% and 10% compared to the original grid. An additional control for spatial data would be to use alternative grids with a spatial resolution identical to the original grid, but slightly shifted vertically and horizontally in space. One may for example generate two grids shifted geographically to the North, South, East, and West by 10% and 5% of the original grid resolution.

Practical Lessons Learned

I: Make Your Case Study Replicable

Each research work is somewhat unique, and this is certainly what makes academic research so fascinating. Selecting appropriate study design, methods, and data require solid knowledge and expertise and cannot be fully automatized; at least for now. Furthermore, one cannot exclude possible dependencies between the results and the selected research design, method and data. As with any academic work, analysis of spatial data needs to comply with current scientific standards. This includes the use of statistical software. For data gathering, spatial statistical analysis and graphics, we used R, a versatile free software that can be used on most operating systems (R Development Core Team, 2011). We used Stata for the modelling fitting processes (StataCorp., 2017). For the sake of replicability, co-authoring, and make research a lot easier in the long run, we highly recommend students to code all procedures (including data gathering and cleaning) into statistical software.

II: Balance between Fineness of the Spatial Resolution versus Computation, Theory, and Data Characteristics

First, there is an intuitive trade-off between the number of dimensions of a problem (number of observations, number of spatial dimensions, parameters of the models etc.) and the computational time required to fit the models. The higher the number of dimensions, the more computationally intensive. Second, the theoretical framework has an impact on the choice of the spatial resolution. When carrying out an empirical work aiming at assessing a theory, one should always ask on what spatial scale(s) does the theory operate. For example, in their desire to provoke civilian population, does one expect terrorist groups to target cities with specific ethnic characteristics—this would require a *local-level* scale analysis—or rather, should the countries in which the attacks are perpetrated matter—the latter would require a *country-level* scale analysis. Third, inevitable spatial inaccuracies in the data often impose huge restrictions on the adopted spatial resolution. This is particularly problematic in

the field of International Relations and Political Science (but not only), where data are often relatively sparse and relatively inaccurate in space. When confronted with sparse spatial data, dividing a study area into too fine squares may inflate the number of zeroes in the statistical models, which would require further treatment afterwards.

III: Creativity

Academic research, including the analysis of spatial data, cannot be carried out without creativity. A proper study design along with a suitable method applied to analyse a comprehensive dataset may not necessarily suffice to address research questions that have not been answered yet. In our paper, recall that we aimed at examining the link between ethnic polarization and terrorism at a local scale, using PRIO-GRID as spatial unit of observation. At the time of authoring our paper, there were no local data available on ethnic polarization. Based on GIS techniques, we developed a method to compute ethnic polarization at PRIO-GRID level, this, for the entire world. The newly-created variable aggregated at a sufficiently-fine spatial scale allowed us to capture the relationship terrorism-ethnic polarization at a spatial level suitable to answer our research question.

Conclusions

Based on a case study, this paper identified key methodological issues that may arise when faced with spatial data. First, we highlighted that the level of analysis used by prior studies was essentially focused at country-level, and hence, failed to capture the effects of ethnic polarization on terrorism. This calls for the use of data gathered at finer spatial resolution and suitable methods to analyse them. Second, we suggested four steps that need to be undertaken for a proper analysis of spatial data. We highlighted the needs for defining clear temporal and spatial scopes before focusing on spatial characteristics of data. We described the underlying assumptions of three common spatial models of data—point, geostatistics, and lattice—and

suggested one approach that allowed us to model our data. We described the processes undertaken to mitigate the issues of encountering excess of zeroes and highlighted the importance of choosing a spatial unit of analysis in line with spatial inaccuracies on the data. We pointed out that the spatial level used to aggregate data has an influence on the characteristics of the dependent variable, whose properties should fit with a chosen probability distribution function.

In his interview given to the *Saturday Evening Post* in 1929, Einstein famously emphasised that "[i]magination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution." As a final word, we encourage the reader to create novel approaches to reveal patterns and mechanisms at spatial scales yet to be discovered.

Exercises and Discussion Questions

Context: You have gathered data that include economic (per Capita GDP), demographic (size of the population), and crime (number of crimes) variables for each of the 26 member states of the Swiss Confederation, also-called "Cantons". Temporal dimension is excluded here; your data represent values averaged over a given period. You are interested in examining the link between the number of crimes (your dependent variable, also-called response) and per Capita GDP (an independent variable) controlling for the size of population (a second independent variable). In line with what you have learned, answer the following questions:

Q1: what is/are possible suitable spatial unit(s) of observations for your analysis? How your choice might affect the interpretability of your results? *[Hint: as stated in the Context, the aim of the research is general enough to allow several alternative spatial units of analysis].*

Q2: what view on your spatial dataset is the most suitable and how do you come to that conclusion? *[Hint: recall that one may consider the process that generate spatial data as points, lattice regular, lattice irregular, or geostatistics].*

Q3: what possible probability distribution(s) could characterise your dependent variable (number of crimes)? *[Hint: recall that the dependent variable is a count of the number crimes, which refers to a count variable. Find information on the Internet about candidate distribution functions adapted to this type of variables and justify their use according to their underlying assumptions].*

Q4: the results of your analysis show that at Canton-level, crime is negatively related with per Capita GDP of the Canton. What modifications can you make in your research to assess the robustness of your results? *[hint: think about all elements included in the modelling process, including the spatial accuracy of the data, the type of your dependent variable, the number of independent(s) variables, the scope of your analysis, etc.]*

Further Readings

Darmofal, D. (2015). Spatial analysis for the social sciences. Cambridge University Press.

DOI: 10.1017/CBO9781139051293.

Web Resources

Zhukov, Y. (2010). Applied Spatial Statistics in R [Slides, R code and datasets]. Retrieved from: <https://scholar.harvard.edu/zhukov/classes/applied-spatial-statistics-r>.

References

Arva, B. J., & Piazza, J. A. (2016). Spatial Distribution of Minority Communities and Terrorism: Domestic Concentration versus Transnational Dispersion. *Defence and*

- Peace Economics*, 27(1), 1–36. <https://doi.org/10.1080/10242694.2015.1055091>
- Choi, S.-W., & Piazza, J. A. (2016). Ethnic Groups , Political Exclusion and Domestic Terrorism. *Defence and Peace Economics*, 27(1), 37–63. <https://doi.org/10.1080/10242694.2014.987579>
- Enders, W., Hoover, G. A., & Sandler, T. (2016). The Changing Nonlinear Relationship between Income and Terrorism. *Journal of Conflict Resolution*, 60(2), 195–225. <https://doi.org/10.1177/0022002714535252>
- English, R. (2010). *Terrorism: How to respond*. New York, NY: Oxford University Press.
- Hoffman, B. (2006). *Inside terrorism* (Revised Ed). New York: Columbia University Press.
- Holt, D., Steel, D. G., & Tranmer, M. (1996). Area homogeneity and the Modifiable Areal Unit Problem. *Geographical Systems*, 3(2–3), 181–200. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-0000586355&partnerID=tZOtx3y1>
- Krueger, A. B., & Laitin, D. (2008). Kto Kogo?: A Cross-Country Study of the Origins and Targets of Terrorism. *Terrorism, Economic Development, and Political Openness*, 148–173. <https://doi.org/10.1017/CBO9780511754388.006>
- National Consortium for the Study of Terrorism and Responses to Terrorism (START). (2017). Overview of the GTD. Retrieved April 7, 2018, from <http://www.start.umd.edu/gtd/about/>
- Nemeth, S. C., Mauslein, J., & Stapley, C. (2014). The Primacy of the Local: Identifying Terrorist Hot Spots Using Geographic Information Systems. *The Journal of Politics*, 76(02), 304–317. <https://doi.org/10.1017/S0022381613001333>
- Python, A., Brandsch, J., & Tskhay, A. (2017). Provoking local ethnic violence – A global study on ethnic polarization and terrorist targeting. *Political Geography*, 58, 77–89. <https://doi.org/10.1016/j.polgeo.2017.02.001>
- Python, A., Illian, J., Jones-Todd, C., & Blangiardo, M. (2018). A Bayesian Approach to Modelling Subnational Spatial Dynamics of Worldwide Non-State Terrorism, 2010 - 2016. *Journal of the Royal Statistical Society Series A: Statistics in Society*, In Press.
- R Development Core Team, R. (2011). *R: A Language and Environment for Statistical Computing*. *R Foundation for Statistical Computing* (Vol. 1). <https://doi.org/10.1007/978-3-540-74686-7>
- Reynal-Querol, M. (1998). *Growth and Religious Conflict*. Mimeo.
- Reynal-Querol, M. (2001). *Ethnic And Religious Conflict , Political Systems and Growth (PhD thesis)*. Retrieved from <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.271096>
- Richardson, L. (2007). *What terrorists want : understanding the enemy, containing the threat*. *Random House trade paperbacks*. Retrieved from <http://ci.nii.ac.jp/ncid/BA85188824>
- START. (2016a). *Global Terrorism Database (GTD) Codebook: Inclusion Criteria and Variables*. Retrieved from <http://www.start.umd.edu/gtd/downloads/Codebook.pdf>

- START. (2016b). GTD CODEBOOK: INCLUSION CRITERIA AND VARIABLES.
Retrieved from <http://www.start.umd.edu/gtd/downloads/Codebook.pdf>
- StataCorp. (2017). No Title. College Station, TX, USA: StataCorp LLC.
- Tanner, M. A., & Wong, W. H. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82(398), 528–540.
<https://doi.org/10.1080/01621459.1987.10478458>
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 234–240. <https://doi.org/10.1126/science.11.277.620>
- Tollefsen, A. F., Strand, H., & Buhaug, H. (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2), 363–374.
<https://doi.org/10.1177/0022343311431287>
- Viereck, G. S. (1929, October 26). What Life Means to Einstein. *The Saturday Evening Post*.
Retrieved from http://www.saturdayeveningpost.com/wp-content/uploads/satevepost/what_life_means_to_einstein.pdf
- Weidmann, N. B., Rød, J. K., & Cederman, L.-E. (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research*, 47(4), 491–499.
<https://doi.org/10.1177/0022343310368352>