11-13-2015

GIS-integrated mathematical modeling of social phenomena at macro- and micro- levels—a multivariate geographically-weighted regression model for identifying locations vulnerable to hosting terrorist safe-houses: France as case study

Elyktra Eisman

Florida International University, elyktra.eisman@ieee.org

DOI: 10.25148/etd.FIDC000154

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

Part of the [Analysis Commons](#), [Applied Statistics Commons](#), [Geographic Information Sciences Commons](#), [Models and Methods Commons](#), [Multivariate Analysis Commons](#), [Other Applied Mathematics Commons](#), [Other Social and Behavioral Sciences Commons](#), [Social Statistics Commons](#), [Statistical Methodology Commons](#), [Statistical Models Commons](#), and the [Urban Studies and Planning Commons](#)

Recommended Citation

Eisman, Elyktra, "GIS-integrated mathematical modeling of social phenomena at macro- and micro- levels—a multivariate geographically-weighted regression model for identifying locations vulnerable to hosting terrorist safe-houses: France as case study" (2015). *FIU Electronic Theses and Dissertations*. 2261.

<https://digitalcommons.fiu.edu/etd/2261>

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

GIS-INTEGRATED MATHEMATICAL MODELING OF SOCIAL PHENOMENA AT
MACRO- AND MICRO- LEVELS—A MULTIVARIATE GEOGRAPHICALLY-WEIGHTED
REGRESSION MODEL FOR IDENTIFYING LOCATIONS VULNERABLE TO HOSTING
TERRORIST SAFE-HOUSES: FRANCE AS CASE STUDY

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

INTERNATIONAL RELATIONS

by

Elyktra Claire Eisman

2015

To: Dean John F. Stack, Jr.
School of Politics and International Relations

This dissertation, written by Elyktra Claire Eisman, and entitled GIS-Integrated Mathematical Modeling of Social Phenomena at Macro- and Micro- Levels—A Multivariate Geographically-Weighted Regression Model for Identifying Locations Vulnerable to Hosting Terrorist Safe-houses: France as Case Study, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Jennifer Gebelein

Mohiaddin Mesbahi

Richard Olson

Thomas A. Breslin, Major Professor

Date of Defense: November 13, 2015

The dissertation of Elyktra Claire Eisman is approved.

Dean John F. Stack, Jr.
School of Politics and International Relations

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2015

ACKNOWLEDGMENTS

I wish to thank the members of my dissertation committee—Drs. Thomas A. Breslin, Jennifer Gebelein, Mohiaddin Mesbahi, and Richard Olson, for their unwavering confidence and encouragement during my dissertation research and throughout my graduate career. This research would not have been possible without the many hours they so generously devoted to guiding me through this daunting task. I also thank them for their infinite patience as I struggled to balance academic work with the demands of a full-time job.

I offer my deep gratitude to my major professor, Dr. Thomas A. Breslin and to Dr. Jennifer Gebelein for their research guidance and mentoring on a host of topics associated with the pursuit of my PhD. I thank them both for graciously accommodating my erratic schedule to convene for sporadic meetings to go over my research. I thank Dr. Breslin for providing insights and historical context that deepened my analysis and maintained my research focus. I also thank Dr. Breslin for recommending journal articles and French news reports directly related to the research that otherwise would have escaped my attention. I especially thank Dr. Gebelein for all the mentoring, research guidance, countless weekend phone meetings, and computer chats to enable her to virtually look over my shoulder at the model iterations.

I also offer my deep gratitude to Dr. Mohiaddin Mesbahi for adding me to a long list of dissertation advisees under his guidance, tightening an already busy schedule by agreeing to join my committee. His guidance as a committee member and as a teacher has been invaluable, enabling me to benefit from the depth of his expertise in Middle Eastern and national security studies. I also thank him for not letting me give up on this dissertation when it was becoming overwhelming, and I could see no light at the end of the tunnel.

I thank Dr. Richard Olson for serving on my committee and for helping me develop the research question and “ways forward” as I embarked on such an intimidating journey. I also thank Dr. Olson for informing my research approach by providing insights into the depth and complexity of variables that can act as societal stressors and ignite crisis events.

I am also grateful to Jeffrey Miller and Joe Gezo for their suggestions relating to the mathematical, methodological, and scientific aspects of this research. Their feedback markedly improved this dissertation—I thank them both for their valuable input.

I thank my parents for always being there, for their love and enthusiastic support for what has seemed like a pursuit that would never end.

Finally, I thank everyone for encouraging me to begin the doctoral journey.

ABSTRACT OF THE DISSERTATION

GIS-INTEGRATED MATHEMATICAL MODELING OF SOCIAL PHENOMENA AT MACRO- AND MICRO- LEVELS—A MULTIVARIATE GEOGRAPHICALLY-WEIGHTED REGRESSION MODEL FOR IDENTIFYING LOCATIONS VULNERABLE TO HOSTING TERRORIST SAFE-HOUSES: FRANCE AS CASE STUDY

by

Elyktra Claire Eisman

Florida International University, 2015

Miami, Florida

Professor Thomas A. Breslin, Major Professor

Adaptability and invisibility are hallmarks of modern terrorism, and keeping pace with its dynamic nature presents a serious challenge for societies throughout the world. Innovations in computer science have incorporated applied mathematics to develop a wide array of predictive models to support the variety of approaches to counterterrorism. Predictive models are usually designed to forecast the location of attacks. Although this may protect individual structures or locations, it does not reduce the threat—it merely changes the target. While predictive models dedicated to events or social relationships receive much attention where the mathematical and social science communities intersect, models dedicated to terrorist locations such as safe-houses (rather than their targets or training sites) are rare and possibly nonexistent. At the time of this research, there were no publically available models designed to predict locations where violent extremists are likely to reside. This research uses France as a case study to present a complex systems model that incorporates multiple quantitative, qualitative and geospatial variables that differ in terms of scale, weight, and type. Though many of these variables are recognized by specialists in security studies, there remains controversy with respect to their relative importance, degree of interaction, and

interdependence. Additionally, some of the variables proposed in this research are not generally recognized as drivers, yet they warrant examination based on their potential role within a complex system. This research tested multiple regression models and determined that geographically-weighted regression analysis produced the most accurate result to accommodate non-stationary coefficient behavior, demonstrating that geographic variables are critical to understanding and predicting the phenomenon of terrorism. This dissertation presents a flexible prototypical model that can be refined and applied to other regions to inform stakeholders such as policy-makers and law enforcement in their efforts to improve national security and enhance quality-of-life.

TABLE OF CONTENTS

CHAPTER	PAGE
I. Introduction	1
II. Methodology	6
Overview.....	6
Sample Data.....	7
Methodology—GIS Platform.....	9
Data for Coefficient Computation—Challenges of incorporating data from multiple sources and in multiple formats	9
France as Test Case	13
III. Related Research	19
IV. Model Development and Design	23
Ordinary Least Squares Linear Regression Analysis	32
Exploratory Regression Analysis to Identify Optimal Model	33
Exploratory Regression Analysis Candidate Variables.....	34
V. Geographically Weighted Regression—Model Variables	41
Household Income.....	42
Education Beyond College	47
Crime	64
Coefficient Behavior in High Local R ² Areas.....	70
VI. Background of suspected extremists—Do They Fit the Model?	74
VII. Factors excluded from the regression analysis.....	77
Social Environment—Impact of Ethnicity.....	77
Societal Membership—and the Political Voice	78
Proximity to Sensitive Urban Zones.....	80
ZUS Components	84
State Policy.....	88
Supportive Population.....	92
Religiosity Coefficients.....	93
Existing Smuggling Routes.....	95
Border Proximity	97
Micro-level Variables	98
VIII. Conclusion	101
Note on Map classification of variables and GWR coefficients.....	110
Bibliography	111
VITA	116

LIST OF FIGURES

FIGURE	PAGE
Figure 1. Scatterplots for Candidate Variables vs. Events	36
Figure 2. Spatial autocorrelation and global Moran's I statistical summaries.....	39
Figure 3. Map Showing Local R Squared Coefficient Strengths	41
Figure 4. Map Showing GWR Coefficient Strength for the Variable, "Household Income", per Department, 2008 INSEE data.	45
Figure 5. Map Showing GWR Coefficient Strength for the Variable, "Education Beyond College", per Department, 2008 INSEE data.....	48
Figure 6. Map Showing Percentage Turkish Ethnicity by Department, 2011 INSEE data.	51
Figure 7. Map Highlighting Departments with Highest Turkish Ethnic Percentages in High Local R Squared Areas.....	52
Figure 8. Map Showing Percentage Moroccan Ethnicity by Department, 2011 INSEE data.	54
Figure 9. Map Highlighting Departments with Highest Moroccan Ethnic Percentages in High Local R Squared Areas, Showing Graduate Degree Coefficient Chloropeth.....	55
Figure 10. Number of Events and Corresponding GINI Indices of Departments.	61
Figure 11. Map Showing GWR Coefficient Strength for the Variable, "Crime" (Burglary Percentages per Population), per Department, 2008 INSEE data.	65
Figure 12. Map Showing GWR Coefficient Strength for the Variable, "Crime" (Burglary Percentages per Population) in a Single Variable Regression Analysis, 2008 INSEE data.	66
Figure 13. Scatterplot Showing Weak Correlation between Burglaries and Events per Department	67
Figure 14. Average Declared Departmental Household Income, INSEE 2008 Data.	72
Figure 15. Number of Events and Percentage Middle-Eastern and North African Municipal Representation in 2008.....	79
Figure 16. Map Showing ZUS Regions and Event Locations.	81

Figure 17. Average Unemployment by Department, 2008 INSEE data.	85
Figure 18. Population Percentage Aged 18-64 with No Education Beyond High School by Department, 2008 INSEE data.	88

I. Introduction

Adaptability and invisibility are hallmarks of modern terrorism, and keeping pace with its dynamic nature presents a serious challenge for societies throughout the world. The rising profile of terrorism has resulted in an acute awareness of both its reach and its destructive capabilities. Enormous resource expenditures are allocated to its eradication, yet the agility with which it responds to countermeasures has endowed it with a lasting resilience. This frontless conflict has required of its targets a level of vigilance exceeding that which many of them have ever experienced. Conventional combat techniques have fallen short of providing a level of security acceptable to the populations they are intended to protect. To that end, both the private and public sectors have experimented with creative technological applications in an effort to understand this enemy with the intention of anticipating its behavior.

Innovations in computer science have incorporated applied mathematics to develop a wide array of models that reflect the variety of approaches to counterterrorism. Many of these are analytical, as it is hoped that added insight will bring more effective solutions, and this is certainly true. Others are predictive. While both types value substantive understanding, predictive models lend their focus to prevention.

Predictive models are generally designed to forecast the location of attacks. Successful prediction is assessed by the model's ability to reveal vulnerable locations so that appropriate defensive measures can be implemented. Although this may protect individual structures or locations, it does not reduce the threat—it merely changes the

target. This reality has not been lost on the security community. Therefore, resources and efforts have been directed at identifying the origin of the threat. Such models frequently employ agent-based modeling and social network analysis, as they focus on the inter-personal relationships that constitute terrorist networks. The value of their contribution is evident from their continued use and from the resources currently devoted to further training and research. While predictive models dedicated to events or social relationships receive much attention where the mathematical and social science communities intersect, models dedicated to terrorist locations such as safe-houses (rather than their targets or training sites) are rare and possibly nonexistent. At the time of this research, there are no publically available models designed to predict locations where violent extremists are likely to reside.

This question requires a complex systems model to incorporate the variety of quantitative, qualitative and geospatial variables that might impact the model. The selection of a safe-house location is dependent on many variables that differ in terms of scale, weight, and type. Though many of these variables are recognized by specialists in security studies, there remains controversy with respect to their relative importance, degree of interaction, and interdependence. Additionally, some of the variables proposed in this research are not generally recognized as drivers, yet they warrant examination based on their potential role within a complex system.

These variables act as constraints and drivers that influence an individual's decision-making process in determining the most viable and safest place to live. Geography is a multi-dimensional variable set that must be included to answer the question. The choice of where to situate oneself has objective and subjective components. The subjective

aspects of decision-making, such as feelings and perceptions about environmental suitability, often have a basis in objectivity. Enclaves comprising a single ethnic or national majority often exist because events encouraged group migrations from one geographic region to another. There was either a driver at the point of origin, a driver to encourage migration to the destination, or a combination of both. The drivers affecting these migrations affect the contemporary social atmosphere within the immigrant communities and the attitudes of citizens of the host country. The location of these enclaves within the spatial context of the host country and the accessibility to and from the country of origin, can play a significant role in how other variables impact the model. For example, an individual might choose to live in an enclave of his home culture because a life among natives of the host nation would expose him to unfriendly attitudes and treatment. Residence within this enclave could restrict access to some resources, just as it could increase access to other resources open only to those in a circle of trust. Residence in the enclave might increase or decrease mobility, depending on where and in what context the individual wanted to travel. Mobility is increased if there is benefit to blending in with a crowd, but can be challenging if one must travel outside the enclave or cover a large distance. Thus, the size and shape of these enclaves and the geographic resources (such as infrastructure) available to them can affect the model. This reality is understood by social scientists, especially geographers, historians, and political scientists. Doreen Massey (2005) characterizes space as being "...about relations, fractures, discontinuities, practices of engagement..." as an "...intrinsic relationality of the spatial..." (Massey, 85). Geography encompasses a set of variables that must be incorporated. Brunsdon et al (1998) presented geographically weighted regression to address the problem of multivariate non-stationarity in geographic space.

For the regression function

$$y_i = \sum X_{ij}\beta_j + \epsilon_i$$

Brunsdon et al estimate $(\beta_j p_j)$ for each variable j and each location i with localized coefficients. They suggest potential localization functions, among them the Gaussian distance-decay-based weighting function

$$\alpha_{ik} = e^{-d_{ik}^2/2h^2}$$

bounding each regression estimate. Here, d_{ik} is the distance between observations i and k . The constant h is a constraint on d_{ik} —the kernel bandwidth—and is described in detail in the chapters “Model Development and Design” and “Exploratory Regression Analysis Candidate Variables” .

GIS platforms can accommodate the plethora of variable types that must be incorporated to conduct this novel approach to identifying where violent extremists might call “home”. GIS platforms can process multiple factors that cannot be adequately modeled with linear analytic technique. Additionally, GIS software environments can present multidimensional models visually in a way that can more easily be understood than by more traditional, textual methods. Additionally, the variables and datasets can be updated and rectified to refine the model, resulting in greater sophistication of analysis. GIS software also has the ability to retain a history of model updates so a user may keep a record of variable changes and model adaptations. This is essential for future

researchers to be able to see the evolution of this model and adapt it for their own uses as well as validate it in future studies.

Most importantly, this dissertation research shows that geography is critical to understanding and predicting the phenomenon of terrorism. Clare Richardson (2011) conducted a multivariate regression analysis examining the power of relative deprivation theory to explain radicalization. Her global study examined 56 countries between 1980 and 2008 to assess the correlation of unemployment and tertiary education levels with terrorist attacks. Since her research analyzes data at the country level, the analysis conceptualizes a country as a community. However, in examining potential drivers of radicalization, this approach limits the depth to which one can study the phenomenon. The country-as-community approach emphasizes shared identity on a macro-level, and assumes that there is indeed a country-level shared identity. The problem of radicalization might stem in part from individuals or small groups not identifying with the host country. In this context, a country-level analysis would overlook that phenomenon.

Central Research Question

This dissertation endeavors to determine the factors that promote an environment that drives and sustains radicalization. Certain assumptions have been made with respect to the nature and identity of these contributing factors in order to focus the research and make it viable. The first assumption is that the most influential factors are social and/or geographical. The second assumption is that the selected factors have the most significant impact in answering the following central research question—

"What factors influence or determine the probability that a region will host a base of operation—or safe-house—for terrorist activities, and if so, why do these factors play significant roles?"

The dependent variable was narrowly constrained to facilitate testing against the set of independent variables. Thus, the dependent variable—the existence of an arrest/raid location—was limited to those raids and arrests that were associated with Islamic extremist ideologies. Broadening the scope to include other extremist ideologies such as militant Basque separatism, would require the inclusion of additional variables and datasets. Some of the variables and datasets might not be relevant to datapoints associated with Islamic extremism. For example, variables that would exacerbate or inhibit Basque extremism might not have any effect on Islamic extremism, as different social groups and value systems are in play.

II. Methodology

Overview

The regression analysis tested multiple variables with and without geographic weighting to determine the existence and importance of geographic factors to the design of a viable model.

This research determined that geographically-weighted regression analysis produced the most accurate model among the regression models tested for this research. Geographically Weighted Regression (GWR) generates a multivariate predictive model that functions in a geospatial information systems (GIS) environment. This model was developed and tested in ESRI's ArcGIS 10.2.

Whereas other types of multivariate regression analysis assume that variables—or regression coefficients—behave uniformly throughout the model, geographically weighted regression (GWR) accommodates non-stationary coefficient behavior. For instance, independent variables might affect the dependent variable in different ways depending on the geographic space, with the coefficient varying locally, rather than having one effect that permeates the model space. For the model in this research, GWR analyzes the coefficient behavior around event locations and outputs coefficient values around these “fit points”.

Sample Data

In this research, the test data are event locations and residences where an arrest or raid occurred and where weapons, bomb-making materials, and extremist literature were discovered by law enforcement. Test data—hereafter referred to as events or event data—were obtained from newspaper reports and online databases from 2008 through 2012. Unfortunately, these sources do not always provide sufficient geographic detail on every event. For instance, an arrest or raid was described as having occurred in Paris, but no other information, such as arrondissement or street address, was provided. This

lack of granularity limits the model's accuracy because there is no localized datapoint around which to test the coefficient behavior. Granularity inconsistencies also caused a clustering effect that artificially increased the dependent variable count in cities that contained multiple event occurrences. These event clusters could not be accurately distributed because the report did not specify where in the city the event occurred. Methodologically, there are different approaches to this problem. Among the simplest and most direct is to eliminate those events from the dataset or distribute the datapoints throughout the city, based on one or more criteria, such as population or even distribution throughout the city's arrondissements. Since there are only 50 datapoints that fit the criteria to qualify as events, or test data, omission was a last resort given that regression techniques, especially those incorporating geospatial attributes, generally speaking, function best with as large a sample size as possible. (Knofczynski & Mundfrom, 2008) (Chang)¹ Distribution of datapoints according to population density was also considered. This method was rejected because its inclusion implies that an independent variable—population density—is a regression coefficient before the model has been generated. Therefore, the clustered events were distributed throughout the city as evenly as possible with respect to arrondissements around the city centroid, not adjusting to any other variable that might artificially incorporate it as a coefficient value.

¹Knofczynski and Mundrom (2008) conducted Monte Carlo simulations to determine minimum sample sizes for different numbers of independent variables and correlation levels. Their analysis indicates a sample size of 21 can produce "good" estimation for an R2value of .7. The GWR in this research yields an R2 of .71.

Methodology—GIS Platform

The analysis uses geographically-weighted regression to examine multiple factors that might correlate with locations where raids and arrests have revealed extremist literature, weapons, or bomb-making materials, or where individuals suspected of planning or participating in terrorism-related activities have resided. ESRI software was selected based on its extensive library of geostatistical tools and its ability to incorporate large datasets, process them and display the results cartographically. The software is also widely available, so the resulting prototype model can be tested and refined by those interested in adapting the model to a different case.

Data for Coefficient Computation—Challenges of incorporating data from multiple sources and in multiple formats

Despite ArcGIS flexibility, model implementation was not a matter of plug-and-play. Datasets were imported in tabular form, but often shape files (a native file format of ArcGIS) had to be built in order to conduct geostatistical analyses and display the data properly. Creating shapefiles is time consuming, labor intensive, and requires a high level of attention to detail when generating georeferenced points for analysis.

Datasets were obtained from INSEE, Eurostat, Data.Gouv.Fr, GADM.org, Pew Research, Diva-GIS, and ESRI. Data granularity, unique location identifiers, and file formats differed and often had to be manually edited before they were integrated in the GIS environment for processing. In addition to manual editing of individual files, one of

the most important factors was table and data integration which allowed data from different sources to be combined and analyzed as an entire data set. The section below details these challenges and the techniques used to surmount them.

The first obstacle was one of semantics. Definitions of communes and arrondissements were often mislabeled, and some sources referred to arrondissements as communes and vice versa. Official terminology is often problematic for performing relational joins. A join is a unique identifier that allows records from two or more tables to be joined together so the entire linked contents can be queried. For instance, a shapefile listed 350 arrondissements, yet there are officially only 342.

The discrepancy might be due in part to the term “arrondissement” being used to refer to two types of administrative divisions. Departmental arrondissements refer to the 342 subdivisions of France’s 342 departments. However, Lyon, Marseille, and Paris are departmental arrondissements that contain municipal arrondissements.² Cantons are a finer level of granularity than departmental arrondissements but are not always strict subdivisions of arrondissements. Cantons contain multiple communes but the boundaries do not always coincide with commune boundaries, as they can straddle communes. INSEE does not list the number of cantons, but its datasets list 3,785 “canton-ou-ville”—canton or city. Other sources list 3,883 cantons or 4,055 cantons³⁴.

²http://self.gutenberg.org/articles/Arrondissements_of_France#References

³Gael L'Hermine. Local Government in France. World Elections Blog.<https://welections.wordpress.com/links/guide-to-the-2014-french-elections/>. List 4,055 cantons.

⁴<http://www.theguardian.com/environment/2011/apr/20/southern-france-sarkozy-shale-gas>. Lists 3,883 cantons.

INSEE defines the commune as the smallest municipal division and lists 36,682. GIS-Diva hosted Database of Global Administrative Areas (GADM) shapefiles consisting of 350 arrondissements. Since the shapefile projected accurately, I used the 350 arrondissements as a guide for obtaining and joining the other datasets.

As a result of these variations in territorial definitions, the independent variable that quantified ethnic representation on the municipal level presented one of the most significant data processing challenges in this research. A dataset that could enable quantification of how well-represented those of North African and Middle Eastern heritage were within their communities could shed light on the impact of having a political voice. I could find no published datasets that quantified this type of representation, largely because France does not gather data on the ethnicity of its political candidates. Data.gouv.fr retains municipal election data, but the downloadable datasets contained voting statistics only—not candidate and winner names. The site does not offer the ability to download datasets that list the names of those elected. To get the data that displays the candidate names, I drilled down to the commune level within each department separately via a multi-step process.

First, I navigated to the page corresponding to a specific department on the website. The department page displays a directory of several pages, each page containing groups of communes in that department. After choosing a page for a group of communes, I selected a link to a commune, to display the data with the winning candidates and the number of votes. There is a download link on each page, but the downloaded data does not contain candidate names. Since there is no way to download the displayed data, I manually copied and pasted directly from the webpage to an excel sheet to save the

data in tabular format. Once on the spreadsheet, I calculated the total number of municipal level representatives by summing several blocks of entries separately because fields are merged. Then, I scrutinized the representatives' first and last names to identify those that were likely Middle Eastern or North African in origin. The percentage ethnic municipal level representation is the ratio of the names I assessed as ethnic to the total number of representatives on each municipal council. This process had to be followed 350 times, resulting in 350 spreadsheets divided among 5 excel files.

I faced similar challenges when gathering statistics on departmental ethnicities. INSEE's data was buried under multiple menu layers. Most of the INSEE datasets only differentiated on the basis of immigrant or non-immigrant, or French or non-French citizen. To discover specific nationalities of immigrants, I located datasets on the regional level and downloaded each individually. Each dataset had to be renamed because all the regions saved their datasets as the same filename. To obtain departmental statistics, I computed percentages from the raw numbers for each department within each region's file. The percentages for each nationality in 101 departments were compiled in a separate table with geographic coordinates of department centroids. Finally, the data could be imported to the GIS software for analysis.

France as Test Case

The enormous datasets common to GIS applications require that this research be limited to a single country case study. France has been selected as the case study primarily for the availability of its social and GIS data. It also contains multiple ethnic communities that have historically presented a security threat, shares borders with several states, and has major sea ports. France has the highest Muslim population percentage in Europe (approximately 7.5%) and tension exists between this population and the host state. Thus, the safe-houses of interest are often those of extremist Islamic actors.⁵⁶

France has a government whose president and senate are elected via its electoral college, enabling its citizens to effect change through the ballot box. Yet, it faces the unwelcome challenge of domestic terrorism. The profiles of these hostile actors complicate analyses, as not all are first-generation immigrants from cultures with a history of religious fundamentalism. Although France offers its citizens free education for those aged 2-16 and a universal healthcare system (Dutton, 2007), many of the hostile actors that threaten France's national security are native-born citizens who were converted and radicalized by Muslim extremists. This reality discredits theories that postulate that repressive governments, lack of educational opportunities, and basic

⁵Population data estimated for 2010. According to Pew Research Center, there were approximately 4.7 million Muslims in France. Germany has the highest Muslim population in Europe—approximately 4.8 million—but that population comprised approximately 5.8% of its population. <http://www.pewresearch.org/fact-tank/2015/01/15/5-facts-about-the-muslim-population-in-europe/>.

⁶2010 Pew Report--"Muslim Population by Country". *The Future of the Global Muslim Population*. Pew Research Center.<http://www.pewforum.org/2011/01/27/table-muslim-population-by-country/>.

services such as healthcare are the drivers of radicalization. The dynamic is not so easily explained.

France's foreign policy is not hostile to states with large Muslim populations. France did not join U.S. forces in the invasion of Iraq—a military action that was unpopular with Middle Eastern states. It supported a multilateral peace agreement concerning Israel and is in favor of the establishment of a Palestinian state (State Department Bureau of European and Eurasian Affairs, 2008).

Therefore, significant factors must be present to account for the severity of France's security problem.

“... France has never fully accepted North African immigrants, and the second generation perhaps even less than the first. That they speak French fluently and readily absorb French culture does not make them welcome in France as earlier waves of immigrants had been...Most North Africans feel they are trapped in a hopeless downward spiral of joblessness, racial discrimination, and clashes with police.” (Tlemçani, 1997)

Although France is officially a secular state, religion plays a significant cultural role. France is historically Catholic,⁷ and the introduction of a relatively recent but significant Muslim population alters its cultural composition. The secular government has actively

⁷ Catholicism was the state religion until 1787, excluding the period between 1592-1685 during which the Edict of Nantes endowed Protestants with limited rights.

and officially supported and promoted its Muslim immigrant communities economically, politically, and culturally. For example, Muslims in military service have access to Muslim clergy and meals consistent with Islamic law. Commercial establishments provide time off and special facilities for Islamic prayer. (Viorst, 1996)

Despite its efforts, France continues to face challenges integrating its immigrant populations, regardless of their ethnicity. The most significant barrier is difficult to pinpoint, but it is likely a combination of factors. Ethnic, religious, and economic drivers are the most frequently cited obstacles for North Africans and Middle Eastern populations.

Doug Ireland (2005), a journalist reporting on politics and culture who spent ten years in France, blames a "...deep-seated, searing, soul-destroying racism that the unemployed and profoundly alienated young of the ghettos face every day of their lives, both from the police, and when trying to find a job or decent housing."

Patrick Simon, et al (2010) encountered racial prejudice in his research on French identity. He refers to France as an assimilationist country because it requires that immigrants identify exclusively as French, explaining that its government and citizens view a hyphenated ethnic identity as incomplete assimilation and thus not genuinely "French". Simon's research analyzes data from the 2008-2009 survey, Trajectories and Origins: Survey on Population Diversity in France (TeO). The Institut National d'Études Démographiques (INED), the Institut National de la Statistique et des Etudes

Économiques (INSEE), and Patrick Simon conducted the research. The survey queried 22,000 residents of mainland France—the largest survey ever conducted that comprised first- and second-generation immigrants as well as non-immigrant French (native-born for two generations). It examined French identity and social cohesion, asking respondents who they consider French, whether or to what degree they perceive themselves as French, and what impacts these perceptions. Respondents were also asked how others perceived their identity, and how their perceptions of others' views affected their self-identification as French. Simon detected a visual component affecting whether or not an individual is accepted as truly French, and that those from predominantly Muslim countries are most negatively affected by that prejudice. Simon observed that “The feeling of having one’s Frenchness denied clearly follows a ‘line of visibility’ affecting primarily immigrants from sub-Saharan Africa and their descendants, then immigrants and descendants from North Africa, Turkey, and Southeast Asia. The pattern is completely different for immigrants from Europe, who feel accepted in the national community.” He adds that people are assessed on whether or not they “look French”. (Simon, 2012) The unspoken requirement to physically appear western European places inescapable restrictions on those who do not conform to that “look”, and the lack of control over how completely one can integrate or assimilate will inevitably breed frustration. The intensity of that frustration is exacerbated or mitigated by many factors both external and internal to the individual, but the ethnic component affecting acceptance as French might play a role in how susceptible one is to extreme ideologies associated with the identity to which that individual is tied.

Adida, et al (2009) differentiated ethnic from religious components by surveying two groups—one Muslim and one Christian—of second- and third-generation Senegalese

immigrants. Both groups are ethno-linguistically equivalent. Adida, Laitin, and Valfort measured the groups' assimilation into French society, defining assimilation as, "... (i) their attachment to their culture of origin; and (ii) their identification with the host culture and society." (Adida et al, 2012) While both groups experienced integration obstacles, the Muslim Senegalese immigrants in the sample were significantly less-integrated than the Christians in the sample. Adida et al's findings are based on 2009 field experiments they conducted on these groups. The experiments consisted of behavioral games designed to measure trust and altruism. Their results showed that native French participants (individuals outside the sample sets whose grandparents are French-born) tend to more often discriminate against the Muslim set because they are more likely to have recognizably Muslim first names. The Muslims in the sample set are more likely to perceive systemic discrimination from French institutions and as a result, choose not to participate in them.

Milton Viorst (1996) argues the barriers have economic roots, but his explanation implies an ethnic component as well, since mostly Northern African ethnic groups responded to the economic driver that generated the integration barrier. Many of France's Muslim youth are descendants of immigrants from France's former colonies. The older generation immigrated to France just after World War II to fill a labor demand to rebuild the city, compensating for the post-war shortage of working-age French males. Many of the immigrants settled permanently, but the labor demand and economic growth did not continue into the coming generations. High unemployment can reduce the motivation to pursue an education if a degree is unlikely to result in a profitable career. The resulting poverty breeds crime, frustration, and fodder for extremists recruiting new adherents.

Religious, ethnic, and economic forces and other conflict points foster an environment that can impede, if not prevent, the integration of Muslim immigrants into French society. Therefore, these individuals may be unlikely to see themselves as Frenchmen even after they obtain citizenship. Their primary identity often rests with either their country of origin or with their religion. Furthermore, perceptions of rejection by the host country (France) and its citizens, combined with a sense of deprivation of opportunities enjoyed by its natives, provide a breeding ground of resentment and hostility that could render them vulnerable to radicalization or to tolerating an environment where extremists can operate.

However, one cannot assume that unwillingness to self-identify as French is the root factor contributing to the development of terrorist safehavens, and that its impact overrides that of an economically vibrant and open society. Terrorists do not establish bases uniformly throughout the country. Nor do they establish them randomly in Muslim enclaves.

Closer examination indicates the presence of multiple contributing factors of varying weights, and these are found in other countries, as well. These factors and their relationships are highly complex, and assume different degrees of importance depending on the country context in which they are considered.

III. Related Research

A body of research has applied quantitative techniques to parse through an array of potential drivers related to some aspect of radicalized behavior, be it criminally or ideologically motivated. Some of these efforts have incorporated geospatial components. Others have focused on cultural or economic drivers. The research presented in this dissertation considers social, political, geospatial, and economic factors. Further, previous research has focused on attacks, pre-attack activities, and training sites. This dissertation research seeks insights into identifying where terrorists might establish their residences. Others have conducted research incorporating geospatial and temporal analytic techniques. The regression model presented in this research does not include temporal analysis due to limitations in the sample set., Suggestions on how temporal analysis might be included in future efforts are included in the text for cases in which an appropriate sample set is available.

Smith et al. (2006) incorporated multiple social and geographic factors with a temporal component to discover timing and distance patterns in preparatory and attack activities associated with terrorism. Smith et al. used datasets provided by law enforcement entities and thus had highly granular data with which to develop and test their research. This data⁸ consisted of 38 data points spanning a 25-year period, each data point consisting of addresses of terrorists' residences, preparation areas, and attack locations

⁸ Some data points did not have sufficient temporal data, so they were not included in the 38 sample points for the aspect of the research that incorporated temporal analysis.

in the United States.⁹ The temporal component of the analysis consisted of dividing the geo-located data into four chronological stages designated by type of activity, from recruitment and preparation to the culminating act, such as a bombing. In reviewing this research, the reader should keep in mind that the study data was limited to the United States, and thus the conclusions of Smith et al. might not apply to terrorist behavior in other countries.

Temporal analysis revealed that the average time period between recruitment and culminating act is relatively short—approximately three months, and that the geospatial pattern is well-defined for international, left-wing, and single-issue (such as animal rights) terrorists. The geospatial pattern is bi-modal, in that nearly half the attacks occurred within 30 miles of the terrorist's residence, but over a third of the attacks were at least 271 miles away.¹⁰

Though the research of Smith et al was geospatially based and incorporated variables such as crime—it included both criminal and non-criminal preparatory behavior—social support systems, its focus was behavioral, and thus did not include the variety of environmental variables that could impact the behaviors and the pace with which those behaviors would unfold. However, they speculated on potential drivers behind the localized spatial pattern associated with international terrorists and in doing so, highlighted variables that could impact extremist behavior. Smith et al suggested that international terrorists' residential locations were restricted by "...new immigration status, lack of transportation, lack of knowledge of the urban landscape, an attempt to minimize

⁹ The types of terrorism included environmental, single-issue, right-wing, left-wing, and international terrorism.

¹⁰Right wing terrorists did not conform to this unusual pattern.

attention or a variety of other reasons.” (Smith, Damphousse, & Roberts, 2006) Variables such as immigration status, unfamiliarity with the urban landscape, and concern about not blending in, allude to several of the variables tested in the regression analyses presented in this dissertation.

Research involving the intersection of terrorism studies with GWR is rare, but innovative research has been conducted using GWR to provide insight on what factors impact a location's vulnerability as a target. Economists Öcal and Yildirim (2013) used geographically weighted regression to study terrorism in Turkey. Their research, while informative, focuses on the behavior of a terrorism index comprised of terrorism-related incidents, injuries and fatalities--not terrorist safe-houses. Dependent variables are primarily economic but do include a schooling ratio and dummy variable for coastal provinces. Analyzing the results of this model design can be challenging, due to the risk of conflating target selection with agent development. To wit, they assess that increases in per capita income might increase a location's attractiveness as a target. This conclusion is logical and borne out in the regression results. They explain that education and employment might be higher in provinces with lower terror indices because individuals who are employed or educated would have more to lose if these areas were hit. This may indeed be true, but this assessment contains an implicit assumption that terrorists live or work in the areas they target. Moreover, the variable choice conflates agency, because the analysis also seeks to explain radicalization drivers as well as incidents, thus conflating two dependent variables. The research presented in this dissertation acknowledges the value of including economic and educational variables, but also considers political, social, and religious drivers. Further, it examines geographic features as candidate independent variables and weighting coefficients to add further

dimension to the GWR model. The inclusion of geographic features as weights or variables in their own right was shown to be not applicable in the test case used for this research, but it could enhance the model in other applications and thus is described in detail in the chapter "Existing Smuggling Routes", "Border Proximity", and "Micro-level Variables".

France's government entities and research institutions collect much of the data necessary to pursue this research. Institut National de la Statistique et des Études Économiques—National Institute for Statistics and Economic Studies (INSEE)—is France's statistical repository. INSEE contains an extensive library of social statistics, some of which are specific to the commune level. Data.Gouv.Fr, another useful data source, is France's open data web presence. It consolidates datasets in csv and xlsx formats from multiple sources, and includes election statistics and data gathered by many of France's ministries. Eurostat provides statistical data to the European Union (EU), and hosts a library of publicly available datasets on EU member states. Pew Research provided data that is the basis for building political representation tables for testing variables, as well as political opinion polls that were sourced for the analysis.

IV. Model Development and Design

The following section explains the development of the regression model, beginning with an examination of candidate algorithms to determine those most appropriate for incorporation in the analysis. Some of the algorithms were developed for counterterrorism applications, such as insurance industry-related terrorism risk models, terrorism attack models for national security applications, and models developed by researchers that identify likely terrorist training sites. Other potential algorithms appear in models developed by the academic community and are not directly related to terrorism. Examples of these include statistical modeling of teenage social structures and alcohol consumption, a kernel weighting function that incorporates geographic space with attribute space to assess tree-stand patterns, and a spectral clustering algorithm to identify criminal communities.

Much of the current research that applies predictive modeling techniques to terrorism concerns risk analysis. The insurance industry is a major patron of such research, as it is aware that physical structures bear varying risks depending on the context in which they are situated. Insurers are also aware that the context is constituted by an interplay of multiple factors at multiple levels. The challenges involved with producing accurate assessments of risk levels require the utilization of complex models with delicately calibrated factors. Yet, successful design and execution of these models is worth the cost of funding these efforts. Insurance companies risk substantial financial losses if a structure is situated in a high risk area, or if the implemented security precautions are inappropriate to the type of risk to which the structure is subjected.

To that end, RAND's Center for Terrorism Risk Management Policy has applied the Probabilistic Terrorism Model designed by Risk Management Solutions to the examination of terrorism in urban settings. The model is often used by the insurance industry to assess liabilities associated with potential terrorist attacks. However, its scope is limited to the United States, and the analyses it generates are based on scenarios involving attack events on fixed targets. Critical factors in the analyses are limited to physical qualities, such as explosive type and building design. It is unlikely that social mechanisms are considered, although this cannot be said for certain, as detailed information regarding this model's design is publicly unavailable. The RAND Corporation was given access to this model for its terrorism research. However, RAND has used it to study discrete event risks, rather than a continuously unfolding situation to which such events can be traced.

Another model designed to serve insurance industry counter-terrorism needs is TerrorRisk, and like the RAND model, it is also focused on potential attack targets. Simon Sole, CEO of Exclusive Analysis, developed this location-based model with Pitney-Bowes MapInfo, and he says it enables his clients "...to precisely differentiate terrorism risk, right down to building level". (Francica, 2007) TerrorRisk uses Pitney Bowes MapInfo to analyze 3,700 locations of interest. The model incorporates pre-assigned numeric coefficients, called "Risk Values", with an insurance company's portfolio to assess its locations' risks of terrorism-related damage. However, information on how it develops the coefficients and what algorithm drives the model is not publically available.

Products such as Risk Management Solutions' Probabilistic Terrorism Model and Exclusive Analysis' TerrorRisk are of a proprietary nature, so their financial value, coupled with the security risk that would result from public release, precludes the possibility of examining these projects in depth. However, some methodologies and their datasets are publically available.

The Open Source Center (OSC) has modeled terrorist attack incidents in Afghanistan and Pakistan using 2004-2008 datasets provided by the National Counter-terrorism Center (NCTC). The geospatial modeling project reveals hot spots that may provide insight into security vulnerabilities and how stakeholders might address regional instability. (Open Source Center, 2009)

The OSC project's predictive modeling component uses geographic features common to most terrorist incidents in the dataset to assess probable locations of future attacks. Its model resulted in a success rate of nearly 69% for locations deemed of high probability. While the OSC's model includes social factors such as roads, population density and ethnicity, this dissertation will consider other factors as well. The incorporation of these additional variables should result in a more accurate and granular model if integrated with sufficient analytical rigor. The proposed model is necessarily more demanding than the Open Source Center's model, as it endeavors to do other than predict an incident. Whereas the primary consideration in attack-planning is the target, with success depending on brief time spans (e.g., the need to access a location just long enough to execute the attack), safe-houses require more lengthy access, reliable ingress and egress, and ideally, opportunities to replenish the human resource pool. Plotting potential safe-house locations calls for greater intricacy of modeling and weighting but

the additional effort can result in significant social benefit, as the identification of at-risk neighborhoods can highlight the need to address issues that will improve the quality of life in those areas and, by extension, the larger region in which they are located. (Open Source Center, 2009)

OSC's work is incident-focused, but counter-terrorism research has also produced location-focused predictive models. This type of research has been conducted by Douglas Way. His work concerns the identification of terrorist locations. (Way, 2003) However, unlike the proposed dissertation research, his work has been focused on locating training camps in remote regions. Training camps differ from safe-houses both in purpose and in context. Therefore, the factors contributing to their establishment differ as well. Additionally, Way's research extends beyond terrorism with religious or political underpinnings, as he is also interested in financially-motivated illicit activities, such as narco-terrorism and arms trafficking. The broader scope of his research necessitates the limitation of his work to geo-physical factors such as terrain and climate. Social variables are generally addressed on superficial levels, consisting of easily quantifiable attributes as state boundaries and population densities. This reduction is necessary, as the increased variation among actors reduces the number of common variables available for consideration.

A variable that Way labels "governance" is the most interesting factor from a social science perspective. He defines this as "...the ability of a nation to control its people." (Way, 2003) Way describes this factor as the most difficult one he assesses in his research. GIS analyses involve applied mathematical techniques and thus require quantitative data. Way assigns a numeric value to each country's "governance" in the

interest of quantifying a contributing factor that is essentially qualitative. However, this appears to be the sole qualitative factor under consideration, and he does not describe his methodology.

In their work entitled *Modeling the Co-evolution of Networks and Behavior*, Snijders, Steglich, and Schweinberger (2007) model alcohol consumption among adolescent friends. They develop a statistical methodology for expressing the relationship between the behavior of individuals and the social structure within which they are positioned. The methodology accommodates actor and social group as joint dependent variables and recognizes the evolutionary nature of this relationship. It is a promising methodological basis for portions of the proposed research because it not only considers actors within their social context, but also considers assimilation patterns. The availability of potential new members is a possible contributing factor to the group's selection of a site location.

Interestingly, techniques developed to assess tree stand patterns may hold even greater potential than those used by behaviorists. In *Incorporating both Geographical Space and Attribute Space into the Kernel Weighting Function of Geographically Weighted Regression*, Haijin Shi, Lianjun Zhang, and Jianguo Liu (2005) present a kernel weighting function incorporating geographical space with what they call "attribute space". This function was developed to study the cluster arrangements of tree species, but it can likely be applied to express spatial arrangements of ethnic group concentrations.

The kernel bandwidth function assigns weights to individual trees based on their distance from a specified point, which can be either a tree cluster point or, for the purposes of this proposed research, an ethnic neighborhood. The function decreases in

weight as the distance increases, so that individuals further away from focal points (or clusters) are less influenced by them. Thus, the spatial weighting function is expressed as the exponential decay of the distance (d) with respect to the kernel bandwidth (h):

$$e^{\left(-\frac{d^2}{h^2}\right)}$$

The kernel bandwidth is the region within which the correlation exists. This region can be defined in terms of geographical space or attribute space (such as income bracket or education). The use of an exponential form is a reasonable starting point for depicting the behavior of geographical space with respect to a selected attribute, as it models the proportional change in related variables. In the context of the proposed research, discrepancies between modeled and actual behavior may indicate that either the relationship does not entail constant proportionality, or that an additional variable influences the actor's behavior. A third possibility concerns the data itself, but this form of equation serves as a reasonable starting point from which to develop a more refined model.

The attribute weighting function incorporates the function $f(\tau)$ which for Shi, et al., represents the difference τ between the focal tree and the surrounding trees. The

resulting equation becomes

$$e^{\left(-\frac{d^2}{h^2} \times f(\tau)\right)}$$

The regression model developed for this dissertation research uses the kernel bandwidth function included in the ESRI ArcGIS environment.

The kernel bandwidth algorithm defines the area around input points within which the coefficient values are determined. For this model, the input points are the geographic coordinates of the event locations. The area constitutes the default search radius, or kernel bandwidth.

The bandwidth is defined by obtaining the mean center of the event points. The distances between the mean centers are measured to obtain the median of all the distances (D_m). These results are the basis for calculating the Standard Distance (SD). The Standard Distance is a single statistic that summarizes the distances that the event locations deviate from the mean center of the event locations.

These values are applied to the following formula:

$$Kernel\ Bandwidth = 0.9 \times \min \left(SD, \sqrt{\frac{1}{\ln(2)}} \times D_m \right) \times n^{-0.2}$$

Where n is the number of event locations.

The identification of current and potential safe-house locations requires an in-depth evaluation of both geographical and social factors on micro- and macro- levels. Demographic datasets from INSEE, Eurostat, Data.gouv.fr, and Pew Research provided the basis for incorporating social and political variables into the model. Geospatial factors such as the types of roads, neighborhoods, and meeting places were considered for inclusion in the model. Though these factors were excluded for a variety of reasons (explained in later sections of this text) from this model, but might well be included were this model applied to a different test case—or country. Thus, the variable sets are contextual. Some variables might be influential in France, but have no effect or be

irrelevant to another country. For instance, local politics might affect daily life far more than national politics if a central government is weak, or if the political structure dictates that administration be more diffuse than centralized. Road density might be inconsequential if the region under study has a robust network throughout.

Gennip, et al (2012) incorporated geographic data with social data to test if the addition of geographic data improved the accuracy of a model they developed to identify street gang communities. Their model was designed for cases with limited geosocial data, and used a spectral clustering algorithm to build network graphs to identify Los Angeles street gang communities.

Their analysis was based on data obtained from LAPD field interviews of 748 suspected and confirmed gang members living in high-density gang areas. Graph edge weighting was based on geosocial similarity between the connecting nodes representing the interview subjects. To increase the accuracy of the clustering algorithm, Gennip et al. included connections within gangs to add ground truth noise and thus increase graph edges (connections that reflect social relationships).

The edge weighting algorithm,

$$w_{i,j} = \alpha S_{i,j} + (1-\alpha)e^{-d_{i,j}^2/\sigma^2}$$

incorporates the Euclidian distance between the average stop locations for nodes i and j. The variable σ is one standard deviation larger than the mean distance between nodes representing individuals who were stopped together. Gennip et al found that the σ value

tended to be constant (1000 feet), enabling the mean distance between i and j to be substituted for σ . They selected the Gaussian distance kernel at the scale to match most social interactions in the dataset and based the design of the social similarity coefficient on the assumption that individuals stopped together share an amicable relationship. Social similarity was expressed using a Boolean matrix where $A_{i,j} = 1$ if $i = j$, meaning that individuals i and j were stopped together. In all other cases $A_{i,j} = 0$. Social and geographic weighting can be incorporated by setting the Boolean α parameter, where zero designates that only geographic information be considered, and a value of one designates exclusive use of social information.

Gennip et al. applied spectral clustering to the graph to obtain 31 clusters, based on LAPD assessment that 31 gangs operate in the area. Their spectral clustering technique applies the k-means algorithm to the first 31 eigenvectors of the normalized affinity matrix to obtain binary approximations to indicator functions for the clusters. The k-means algorithm develops the clusters by iteratively grouping the nodes with their nearest centroids.

According to Gennip et al., in the gang communities studied, the geographic stability of gang territories enhanced the accuracy of the analysis, even though data was sparse and connections could extend beyond the policing area. These characteristics are shared by France's northern African and middle-eastern immigrant communities, as they developed over generations and are thus geographically stable, and their communities do not conform strictly to policing districts.

Data was too coarse to incorporate this spectral clustering technique into the dissertation model. Higher data granularity, such as addresses of event locations and addresses associated with suspects' residences—could identify other extremists and indicate other potential safe-house locations if network analysis could be integrated with the geographically weighted regression model. This could be achieved by incorporating metrics generated by the algorithm as regression variables in the GWR model.

Ordinary Least Squares Linear Regression Analysis

Model testing began with Exploratory Ordinary Least Squares (OLS) Regression to obtain the OLS model with the best fit without considering the impact of geography. The OLS regression model using the variable combination with the closest fit was compared with a Geographically Weighted Regression (GWR) model that used the same variable set. Comparison of the models demonstrated that the most accurate regression model incorporated geographic weighting and provided the basis for the analysis of the regression results.

Ordinary least squares (OLS) analysis—Linear Regression analysis—uses least squares to determine how a set of independent variables can best predict the behavior of a dependent variable.

OLS produces a linear equation that incorporates the independent variable coefficients and returns dependent variable values. The returned values are compared with the dataset of observed dependent variable values to assess the accuracy—or fit—of the

regression model by squaring and averaging the distances from the data points to the line. The least squared value—shortest distance from the line—is used to construct the regression line. This technique is limited, as it produces a linear model and assumes constant variance among the error estimates for each measure. It does not incorporate temporal, or spatial variations of the datapoints. This type of regression analysis was conducted before the geographically weighted regression analysis, firstly, to evaluate which candidate variables should be included in the model and, secondly, to compare the results of both regression techniques to assess the relevance of including a spatial component in the analysis.

Exploratory Regression Analysis to Identify Optimal Model

Exploratory regression analysis is a necessary precursor to conducting OLS analysis when the behavior and impact of multiple candidate variables on an independent variable is not well understood. Exploratory OLS tests multiple candidate variables to determine which combination of coefficients produces an equation whose dependent variable output most closely resembles the observed data. This technique also identifies multicollinearity that must be eliminated to accurately evaluate each variable's contribution to the model. Multicollinearity—correlation between candidate coefficients—prevents successful assessment of each variable's influence on the dependent variable.

Exploratory Regression Analysis Candidate Variables

Data and time constraints limited the choice and number of variables under consideration. Variable selection was based on social science literature, dataset availability, and the goal of analyzing research results in a timely manner. Candidate variables that were considered but not included in the exploratory regression are discussed in the section, “Factors Excluded from the Regression Analysis”. Variables were excluded for a variety of reasons that are explained in depth in that section and that generally relate to the unavailability of a dataset or an analytic assessment that, while a variable might affect candidacy of an area as a safe-house location, that variable is not applicable to France.

Central to this realization is the caveat that the geographically weighted regression model presented in this research is not a global model. One should not simply insert another country's equivalent datasets and expect to generate a regression model with a good fit, even if that model is geographically weighted. Variations in government structure, the catalogue of factors that comprise the social environment, and characteristics of geographic terrain that shape the geostrategic and social environments, can necessitate the incorporation of variables that are not applicable to France's situation. Every region is unique in this sense. The variables excluded in this model might be critical to a model designed for a different region. In quantitative models of social science phenomena, especially those with a geospatial component, dataset exclusions are often related to the data itself. The dataset might not be designed such that it can be incorporated into a model. Different scales, formats, date ranges, and levels of granularity represent a small sample of the obstacles that a social scientist or

geospatial analyst might encounter in the course of his or her research. The methodology however, is widely applicable, and in this sense the geographically weighted multi-variate regression model presented here can serve as a foundation for the design of similar models for other geographic regions.

The variables included in the exploratory regression are the percentage of graduate degree holders per departmental arrondissement in 2008, GINI standard of living index per department in 2008, percentage of Algerian-Moroccan-Tunisian-Turkish immigrants in 2011, percentage ethnic North African and middle eastern representation in municipal government in 2008, burglaries per department in 2008, presence of one or more Sensitive Urban Zones (ZUS) per arrondissement,¹¹ percentage of unemployed per department in 2008, percentage of population aged 15-64 with only primary school education by departmental arrondissement in 2008, and declared taxable revenue per departmental arrondissement in 2008.

The graphs below show correlations between each candidate variable considered in the exploratory regression models and arrest/raid events.

¹¹ Please see section, “Proximity to Sensitive Urban Zones” for explanation of ZUS.

Figure 1. Scatterplots for Candidate Variables vs. Events¹²

¹²“Event” is a datapoint that references the residence of an arrested terrorist suspect or the location of a raid site.

Table 1. Summary of OLS Results

Variable	Coefficient [a]	Std. Error	t-Statistic	Probability [b]	Robust SE	Robust T	Robust Pr [b]	Variance Inflation Factor [c]
Intercept	0.002409	0.033344	0.072243	0.942436	0.035566	0.06773	0.946026	-----
Graduates	-0.016031	0.003202	-5.00726	0.000001*	0.005089	-3.15029	0.001785*	1.594625
Burglaries	0.000033	0.00001	3.396878	0.000774*	0.00002	1.66080	0.09767	1.642839
Income	0.000019	0.000001	20.36246	0.000000*	0.000002	8.07699	0.000000*	1.345643

Ordinary Least Squares Diagnostics

Table 2. OLS Diagnostic Statistics Summary

Input Features:	Observed Events
Number of Observations:	350
Multiple R-Squared [d]:	0.613722
Joint F-Statistic [e]:	183.242935
Joint Wald Statistic [e]:	87.004074
Koenker (BP) Statistic [f]	53.176187
Jarque-Bera Statistic [g]:	2860.941648
Dependent Variable:	EVENTS
Akaike's Information Criterion (AICc) [d]:	243.804925
Adjusted R-Squared [d]:	0.610373
Model Variables	-PRCGRADDEG +INCOME_2008 +BURG2008

Testing Auto-Correlation for OLS Multivariable Regression Analysis

A Morans I spatial correlation test was performed against the multi-variable OLS model prior to conducting geographically weighted regression to evaluate clustering distribution of deviating predictions. If the error terms are randomly distributed, the variables are

likely sufficiently explaining the observed data, and it is unlikely that a crucial variable is omitted from the model.

Selected Parameters:

Input field: Standard Residual

Conceptualization of Spatial Relationships: INVERSE_DISTANCE

I selected the inverse distance method because social phenomena do not behave discretely--nearby features and phenomena have a mutual influence. It is a safe assessment that for this research, the distant features have a reduced impact on each other. The inverse distance method estimates a point's predicted attribute value from the weighted average of the values of proximate points. The value weights are inversely related to the distance between the prediction and the samples, and are modified by a constant power parameter p . As p increases, the distant weights decrease and reach zero, and at very high p values only the nearest points weight the prediction. If $p = 0$, the prediction will be the mean of all the neighboring data values. The inverse distance squared weighted interpolation sets $p = 2$. It is ArcGIS's default value and was the selected parameter for this analysis.

Distance method: Euclidean

Standardization: None

The results appear below:

Figure 2. Spatial autocorrelation and global Moran's I statistical summaries

Geographically weighted regression was conducted to determine if there might be a spatial component that would cause an independent variable's influence to change depending on location. Below is an explanation of the parameters selected for the GWR analysis.

Bandwidth and Kernel type

AICc (Akaike Information Criterion) Bandwidth: For this analysis, AICc was used to facilitate analysis of multiple regression models. AICc estimates the information loss that results from using a given model to predict the behavior of the observed data. It is purely a comparative metric, so there is no typical value range for any field of research or type

of regression model. The lowest AICc value among the candidate models reflects the model with the least information loss of all the models under consideration.

The AICc value can be used to determine the optimal number of neighbors for the estimations. For this analysis, AICc bandwidth units were determined using an adaptive Kernel setting, since feature class density varied throughout the geographic space.

GWR and OLS results appear below:

Table 3. Comparison of OLS and GWR statistics

Measure	OLS	GWR
Neighbors		64
Residual Squares		24.287611477887584
Effective Number		63.974417936184381
Sigma (estimated standard deviation of the residuals)		0.29140027981237099
AICc	243.804925	178.58006025877796
R ²	0.613722	0.76514907676923771
R ² Adjusted	0.610373	0.71344181308492494

The OLS model had an AICc value of 243.805, whereas the GWR model had an AIC of 178.580, an indication that the geographically weighted regression model would be a better fit than the ordinary least squares model.

The OLS model's adjusted R-squared value is 0.61, lower than the adjusted R-squared value of .71 in the GWR model.

V. Geographically Weighted Regression—Model Variables

The geographically weighted regression analysis incorporates three variables as model coefficients—income, burglary and graduate degree holders. The income coefficient is based on the total declared household income per department in 2008. The burglary coefficient is based on the percentage of burglaries per overall population aged 15-64 per department in 2008. The graduate degree coefficient is based on the percentage of graduate degree holders per population aged 15-64 per department in 2008.

Figure 3. Map Showing Local R Squared Coefficient Strengths

The local R squared map shows how closely the model adheres to the observed data—the darker the color, the more closely the model fits the observed events. The lightest shade—corresponding to a local R squared value below .50—reflects those areas where the model predictions are no better than chance occurrences. As shown in the map, the model accuracy varies spatially. For this reason, geographically weighted regression analysis was favored over ordinary least squares regression analysis, which does not consider spatial variability.¹³ A comparison of model statistics provides quantitative evidence that GWR produces a more accurate model.

The model fits best (71 to 87%) in the southeastern and southwestern regions. Roughly speaking, the regions are analogous to Provence Alpes-Cote-d’Azur, Midi-Pyrenees, Languedoc-Roussillon, and Aquitaine, and south-western Bourgogne in central-east. The region with the weakest fit is a north central to north eastern region that stretches down to the center west to Poitou-Charente and Limousin. Later sections will examine why this region is a poor fit.

Household Income

The income coefficient reflects the total taxable household income per departmental arrondissement in 2008.

A body of academic research concludes that extremists tend to be located in poorer areas or areas where there is less policing which also tend to be lower income.

¹³ Rejected OLS and GWR models appear in the appendix.

Researchers such as Townsend, Niebuhr, and Fields, examine the influence of individual factors on radicalization. Proponents of theories that low education, low opportunity, and low income correlate with radicalization tend to deconstruct one or more terrorist activities and equate aspects of the extremism with other anti-social behaviors. For example, the violence and illicit fund-raising activities associated with terrorism are seen as criminal acts.

The tendency to lump terrorist activities with “conventional” crime is fueled by revelations that terrorists and organized crime groups sometimes cross paths and even work together. Garry Reid, deputy assistant secretary of defense for special operations and combating terrorism informed the Senate Armed Services Committee’s emerging threats and capabilities subcommittee that terrorists are employing criminal tactics to finance their operations and in some cases, have formally cooperated to achieve a common end. (Parrish, 2012)

Though there is sometimes a functional overlap, designating terrorism as a criminal type can hinder analysis, primarily because extremists and criminals do not share the same motivations. Thus, crime theory alone is not an adequate mechanism with which to understand radicalization.

Other theorists highlight the apparent self-destructive, high-risk behavior associated with terrorism—especially suicide terrorism—and apply psychological theories that characterize terrorism as desperate actions perpetrated by those who believe they have nothing to live for. Ellen Townsend (Townsend, 2007) asks if terrorists are suicidal. Niebuhr (1960) notes that politically-motivated self-immolators commit suicide to elicit sympathy from others, and do not aim to kill others in the process. Individuals who

commit these acts generally object to being characterized as performing a suicidal act. They view their action as *ijtihad*, sacrificing their lives to Allah—not as self-immolation. According to Fields, et al (2002) (Fields, Elbedour, & Hein, 2002)¹⁴ interviews of families of suicide attackers, most family members characterized them as deeply religious. This is not generally a quality shared by others who self-immolate under emotional depression. In fact, empirical research reveals that religiosity often prevents one from committing suicide, because most religions forbid it.¹⁵ While many of those who attempt suicide cite a desire to escape a painful situation, the majority of suicide attackers cite harm visited upon them or their relatives as a motivation.¹⁶

In addition to the analytic errors associated with attempts to understand radicalization within inappropriate frameworks, deconstruction itself is a risky analytic technique when applied to social science topics. Integrative approaches can offer more insights because they recognize that human behavior is driven by a complex interplay of multimodal variables, psychological, historical, and geospatial. Geographically weighted regression analysis is one such approach, well-designed interview studies are another.

GWR results indicate that income is positively correlated with events in areas where the model shows the strongest predictive power (See map below). This discovery is consistent with the findings of other empirical research studies.

¹⁴ Cited in Townsend.

¹⁵Dervic et al. 2004. From Ellen Townsend.

¹⁶Kushner, 1996.And Bancroft et al. 1979. Both in Ellen Townsend.

Geographically Weighted Regression Coefficients Declared Departmental Household Income 2008

Figure 4. Map Showing GWR Coefficient Strength for the Variable, “Household Income”, per Department, 2008 INSEE data.

Research on the correlation of wealth with violent extremism (specifically, suicide bombers) from the Palestinian territories identified positive links. Berrebi (2003), found that Palestinian suicide bombers were rarely from impoverished households. According to Berrebi, who studied the biographies of 335 Palestinian terrorists, “...those with higher education and higher living standards are more likely to participate in terrorist activity...” (Berrebi, 2007, 2)

Shafiq and Sinno (2009) examine correlations of income and education—two of the geographically weighted regression model coefficients—with an individual's support for suicide bombing attacks on civilians in their countries, and on Western military and political personnel in Iraq. Their research is based on 2005 public opinion data from residents of Indonesia, Jordan, Lebanon, Morocco, Pakistan and Turkey. Comparison of their findings is limited to support for attacks on civilians in their home countries. This is because support for these attacks might vary due to the attacks occurring far from home and are unlikely to affect the respondents or their immediate social groups in the latter case. For France, the context of ethnic group support of extremists storing weapons, recruiting, and living in their midst more closely resembles the context in which respondents are queried on their support of suicide attackers on civilians in their residential (home country) environment. Comparison is also limited to two countries—Turkey and Morocco—as ethnicity data for French departments was unavailable for other ethnic groups in Shafiq and Sinno's study. Shafiq and Sinno found no significant correlation between income and support for suicide bombing. The areas where the GWR's strongest predictive ability coincides with the highest proportion of Turkish and Moroccan populations are Bretagne, Languedoc Roussillon, and western Provence Alps Cote d'Azur.

One potential explanation for the observation that safe-houses are more likely to be found in wealthier areas is that police might put more energy into fostering positive relationships with members of the communities, resulting in a greater willingness on the part of residents to support the police. This would result in a better network of informants who would notify the police whenever they have reason to believe that extremists are stashing weapons or recruiting members. This explanation would also fit the behavior of

the burglary coefficient described in the following section, and could be relevant to other countries that show similar coefficient behavior-event correlation. However, this explanation does not appear to apply to France, as the section, “Police-Community Relationships” details research that counters this conclusion.

Education Beyond College

Though educational level is a predictor in the regression model, the nature of its correlation—most notably in high local R^2 areas in the south and northeastern areas of France—varies with respect to the geographic region (see map below). This variable more than any other, demonstrates the importance of incorporating a geographic component into the analysis.

Geographically Weighted Regression Coefficients Education Beyond College 2008

Figure 5. Map Showing GWR Coefficient Strength for the Variable, "Education Beyond College", per Department, 2008 INSEE data.

The impact of geography as a determinant is evident from other empirical research, as well. Analysis of social, political and historical context provides insight as to why a coefficient has a positive or negative influence for a particular ethnic group. Sinno and Shafiq talk about cultural issues associated with extremist sentiment.

Shafiq and Sinno's research on suicide attackers reveals variations in education, depending on the attacker's country of origin. They find that any correlation between education and support for suicide attacks is difficult to detect, as those with limited

education were less likely to express their opinion, opting instead to select “don’t know/Refused” in answering the questions.

Among those who did answer in Morocco and Turkey, both of which are countries of origin for many Muslim immigrants to France, Shafiq and Sinno found no significant correlation between education and support for suicide attacks on civilians. GWR analysis shows that in departments having the highest proportion of ethnic Turkish and Moroccan populations, the education coefficient has the weakest determining influence. Thus, the behavior of the GWR and the behavior of its education and income coefficients is consistent with Shafiq and Sinno’s observation that education and income are weak predictors of support for extremism among Turkish and Moroccan populations.¹⁷

In France, Turkish departmental population percentages range from 0.1 percent and 23.4 percent of the immigrant population. However, Turkish populations tend to be concentrated in areas where the model has the weakest fit. A comparatively high Turkish population—between 8.1 and 23.4 percent—exists in some areas of the GWR model that have a moderately high local R²—between .51 and .70—somewhat high for a social science regression model. These upper range R² areas are: Bretagne (Morbihan), Pays de la Loire (Maine-et-Loire), and Rhone-Alpes (Drome). Higher education is negatively or positively correlated, or shows no significant correlation, depending on department. In

¹⁷At time of this research, I was unable to obtain data on France’s ethnic percentages for other ethnic groups studied by Shafiq and Sinno.

departments where a correlation is present, in Morbihan it is positively correlated, but in Maine-et-Loire and Drome it is negatively correlated.¹⁸

Thus, where the model results demonstrate the best fit, there is no significant correlation between highly educated Turkish populations and event locations—consistent with Shafiq’s and Sinno’s observation of no significant correlation of highly educated residents in Turkey with support for suicide bombers.

¹⁸In the Auvergne region the higher education coefficient is neutral in the north and weakly positive in the south.

Percentage Turkish Ethnicity by Department

Figure 6. Map Showing Percentage Turkish Ethnicity by Department, 2011 INSEE data.

High Percentage Turkish Departments in High R² Areas with Graduate Degree Coefficient

Turkish representation between 8.1% and 23.4% in areas with local R² between .51 and .70.
Shown with choropleth map of GWR coefficients for graduate degree correlation.

Figure 7. Map Highlighting Departments with Highest Turkish Ethnic Percentages in High Local R Squared Areas.

France's Moroccan departmental population percentages range from 1.9 percent to 42.2 percent of all immigrants. A comparatively high percentage of ethnic Moroccans—between 14.9 percent and 42.2 percent—exist in areas where the local R² is between .70 and .87—the upper range of the regression model's local R², and thus where it demonstrates the best fit. These upper range R² areas are: Aquitaine (Lot-et-Garonne, Gironde), Midi-Pyrenees (Tarn-et-Garonne, Tarn), Auvergne (Haute Loire), Languedoc Roussillon (Gard, Aude, Herault), Provence-Alpes-Cote d'Azur (Vaucluse, Bouche du Rhone, Var), Rhone Alpes (Drome). Higher education correlation in the Moroccan areas

is similar to that of the Turkish areas in that it varies widely. It is positive, negative, or shows no significant correlation, depending on the department. In departments where a correlation is present, higher education is positively correlated with the probability of an arrest/raid location in Aquitaine (Lot-et-Garonne), Midi-Pyrenees (Tarn-et-Garonne, Tarn) and Auvergne (Haute Loire), but negatively correlated in Languedoc Roussillon (Gard, Herault), Provence-Alpes-Cote d'Azur (Vaucluse, Bouche du Rhone, Var), and Rhone Alpes (Drome). There was no correlation in Aquitaine (Gironde) and Languedoc Roussillon (Aude).

Percentage Moroccan Ethnicity by Department

Figure 8. Map Showing Percentage Moroccan Ethnicity by Department, 2011 INSEE data.

This erratic correlation range for areas with higher Turkish and Moroccan concentration explains the exclusion of ethnic concentration as a model variable. According to the regression model results, ethnicity—at least among the ethnic groups for which data could be included in this research—is not a determining factor for the establishment of a safe-house location. These results are consistent with Shafiq and Sinno’s conclusions that there is no significant correlation between higher educational attainment with support for suicide bombers among Moroccan residents who were surveyed.

High Percentage Moroccan Departments in High R² Areas with Graduate Degree Coefficient

Moroccan representation between 14.9% and 42.2% in areas with local R² between .70 and .87. Shown with chloropeth map of GWR coefficients for graduate degree correlation.

Figure 9. Map Highlighting Departments with Highest Moroccan Ethnic Percentages in High Local R Squared Areas, Showing Graduate Degree Coefficient Chloropeth.

Why does higher education's predictive value vary according to ethnicity?

Any attempt to answer this question requires an in-depth investigation of the context in which these ethnic groups develop, and may or may not be linked to the quality of education or the enhanced income opportunities for graduate degree holders in that environment. Shafiq and Sinno explain that in Turkey, higher education correlates with secular views.

There is usually a negative correlation between the percentages of the population aged 15 to 29 years old with education degrees beyond that of college and the presence of events.¹⁹ Areas where the correlation is positive are generally areas where the overall model is a weak predictor. Taking the model at face value, educated communities discourage the establishment of extremist safe-houses. But what mechanism associated with higher education exerts this effect?

Professor Kamaldeep Bhui's (2014) research team at Queen Mary University of London interviewed 600 Pakistanis, Bangladeshi, and Muslim people in London and Bradford between the ages of 18 and 45. The study found no evidence substantiating a link between inequality-related grievances and radicalization. Further, Professor Professor Kamaldeep Bhui (Bhui, 2014) stated to al-Jazeera that his team's research did not reveal, "...what factors make potential recruits open to persuasion to join a terrorist movement...". (Pizzi, 2014) Similarly, further research is required before one can offer insights into why and how advanced education could contribute to radicalization. Any attempts to shed light on this link necessitate speculation, but the GWR results might indicate a starting point for such research.

¹⁹ I used 29 as the cut-off year because the next age bracket included everyone over the age of 30. I was concerned that inclusion of the over-70 population might skew the data. The over-70 population includes generations where undergraduate degrees might have been valued differently, be it changing educational standards or different degree requirements necessary for gaining desirable employment.

Answering the Why—Income and Education Through the Lens of GINI

A possible explanation for the positive influence of higher education on the establishment of extremist safe-houses might be found from examining the GINI Standard of Living Index. The GINI Standard of Living Index measures levels of inequality among living standards. It does not measure the quality of the living standard itself, but rather the disparity of living standards within a specific area. For example, a GINI living index of zero reflects a perfect equality among residents, but does not indicate whether or not that living standard is high or low.

Relative Deprivation Theory posits a positive correlation between high education and violent behaviors (such as crime and terrorism). A logical extension of the theory is that terrorism tends to develop in areas where there is a stark contrast in living standards, because those less fortunate are regularly exposed to the trappings of the higher class, and thus there are constant reminders that they are less well-off. Ted Robert Gurr (1970) writes of this theory in *Why Men Rebel*, in which he explains that the gap between expected and achieved welfare generates violence born of frustration when individuals see no way to attain the lifestyle they feel they deserve. One would expect this frustration to be more acute if an individual follows societal rules and acquires an advanced degree, only to discover additional obstacles that academic achievement will not remove. Thus, proponents of this theory might argue that areas are more at risk to hosting a safe-house if there is quality of life disparity, or where those who enjoy a high quality of life live near or among those less fortunate.

Data note on educational attainment: To address that hypothesis, data on the percentage of higher education graduates in 2008 were analyzed with data on 2011 executive and professional employment. The reasoning for the time gap is to allow three years for graduates to obtain and be promoted into executive and professional positions commensurate with their education and skill level. Regression with these variables produced a Durbin-Watson statistic of 2.24. The Durbin-Watson statistic measures the correlation between adjacent error terms and returns values between 0 and 4. A Durbin-Watson statistic close to zero indicates a positive autocorrelation. If the statistic is close to 4 the autocorrelation is negative. In this instance, the Durbin-Watson statistic was very close to 2, indicating the absence of autocorrelation between those with graduate degrees and those holding professional or executive positions.

To gain insight into whether or not Relative Deprivation Theory can explain event presence, departments with a high GINI index were examined. Income and higher education coefficients were examined for arrondissements within high GINI departments that had at least one event. Results supporting this theory would show events in low income and low graduate degree areas. In such cases, one explanation would be that the living standard disparity is a source of frustration and resentment for those with less education who reason that they could not improve their standard of living because they lack the education necessary to secure a white-collar position. A different result that would also support the theory is the presence of events in arrondissements with low income and high graduate degree coefficients. In such cases, the explanation consistent with the theory asserts that graduate degree holders in low income areas are vulnerable to engaging in anti-social behavior. They have followed societal rules and deferred rewards to meet the challenges associated with obtaining a higher education with the

intent of bettering their lifestyle. When the graduate degree ceases to show promise of opening the door to a profitable career they become disillusioned and anti-social.

GINI indices ranged from 21.13 for Aurillac, Mauriac, and Saint-Flour in Cantal, through 36.08 for Nanterre in Hauts-de-Seine and 42.23 for Paris. Departments in the GINI index brackets over 28 contained 33 out of the 50 events in this study. In descending order and with events in parentheses, they are Paris (9), Bouches-du-Rhône (6), Haute-Garonne (4), Alpes-Maritimes (3), Hérault (2), Rhône (2), Seine-Saint-Denis (2), Var (1), and Vaucluse (1). These events were not concentrated in low income arrondissements, but the distribution does not necessarily eliminate relative deprivation as a driver. In areas such as Paris and Lyon (in Rhône), the GINI index might be high within—not just among—arrondissements, as high income neighborhoods can abut low income neighborhoods. The news reports that serve as the source of the event data rarely specified the address of the raid or arrest location, preventing confident assessments as to the proportion of events in low income neighborhoods. However, in reports that did specify the neighborhood where the events occurred, biographic information in the news articles appears to support relative deprivation theory, if only in the context that certain raids occurred in low income areas.

For example, in 2002²⁰ Mourad Benchellali, his brother Menad, and his father Chellali Benchellali (a Muslim cleric) were arrested in 2002 on terrorism charges. The Benchellalis were residents of Venissieux, a low income suburb in Lyon. These arrests are not included in the research data because they pre-date the study's time window that

²⁰ Although 2002 falls outside the date range of the regression data, it is temporally proximal and thus likely reflects the context of the events within the research data.

scopes the analysis to events occurring between 2008 and 2012. However, the 2002 arrest provides weight to the argument that the Lyon events appearing in this study might well have occurred in low income neighborhoods. This is especially true given the recent arrests in 2014, when two teenage girls—one a resident of Venissieux—were arrested for planning to bomb the Great Synagogue of Lyon.²¹

Another example is Marseille (in Bouches-du- Rhône). Marseille appears as a high income, high GINI area, when in reality it contains both high and low income neighborhoods. Two arrests in 5 October 2010 reportedly occurred in southeastern Marseille, but even this is not sufficient granularity to determine whether or not the suspects lived in a low income neighborhood.

Since nearly 2/3 of the events occurred in high GINI index departments, a high GINI index would appear to be an influential variable but it never appeared as a coefficient in the strongest models over multiple iterations of OLS and GWR analyses. GINI could actually be a weaker determinant than income or crime because while most events occur in high GINI areas, enough events appear in very low GINI areas to weaken the influence of the index as a coefficient, and some high GINI departments do not have event occurrences, though this cannot be said without caveat based on the small sample size and the narrow time frame of analysis. It is possible that relative deprivation is a driver only when other factors are in play, such as high education coupled with high income. However, lack of specific information on the locations of many of the events prevents a confident assessment of this variable's impact.

²¹ The other girl resided in Tarbes, Hautes-Pyrenees.

It is possible that in France, lower income neighborhoods in high GINI areas are vulnerable to developing safehavens for advocates of violent extremism. This could be true for other nations, but the impact of the coefficient combination of high education, high income, somewhat lower crime coefficient, and high GINI index might not be as potent if other variables have sufficient strength to significantly impact the model.

The graph below shows the correlation of events with a combined income-percentage graduate degree holders coefficient for arrondissements and other district types within departments having GINI indices of 28 or higher. Paris and Marseille data are excluded from the graph due to the lack of location granularity. Most of the news articles reporting Paris and Marseille raids did not specify the arrondissement, and both cities contain many districts of widely varying income levels.

Figure 10. Number of Events and Corresponding GINI Indices of Departments.

The coefficient behavior is erratic for GINI indices below 36, but beyond 36, the trend generally increases.

Impact of High Education and High Income—Getting to the Why

A 2014 Queen Mary University of London research study links wealth, full-time education, and youth with radicalization and in so doing, corroborates the geographically weighted regression research in this dissertation. Professor Kamaldeep Bhui's (Bhui, 2014) research team interviewed 600 Pakistanis, Bangladeshi, and Muslim people in London and Bradford between the ages of 18 and 45. Sympathy for violent ideologically-motivated protests and terrorist activity was highest among those under 20 who were unemployed full-time students. Sympathy levels were also highest in the sample portion that earned over £75,000 per year. The study identified those in the pre-radicalization stage, a point in the radicalization process where intervention measures are likely to prevent an individual's transition to a rigid extremist mindset.²²

Nasra Hassan (Hassan, 2001) interviewed approximately 250 Palestinian extremists from 1996 to 1999. All the individuals interviewed either had attempted suicide missions, their family members, or mission trainers. According to Hassan, none of the would-be suicide attackers were poor or uneducated. Most were middle class, employed, and in the case of two volunteers, children of millionaires. Hanadi Jaradat came from a well-to-do Palestinian family and graduated law school before dying via suicide attack on a Haifa restaurant in 2003 (Victor, *Raising a Generation of Martyrs*, 2003).

The “why” is challenging to address. Even with additional, in-depth, interview-based studies motivation is difficult to determine, given that subjects might not be forthcoming

²² Kamaldeep Bhui, Professor of Cultural Psychiatry & Epidemiology, Queen Mary University London.

in their responses, and their root motivations might not be clear to themselves. Azam (2010) suggests that higher education often requires young people to move far from their families, and they become vulnerable to social pressure as they strive to replace loosened family ties with other social connections that are familial in nature. These young, highly educated people might be vulnerable to stronger personalities who encourage them to replace links to their biological families with those of a new, extremist family.

Marc Sageman (2004) reaches a similar conclusion, based on interviews and biographical studies of 172 jihadists. Although a majority (approximately 60%) of his subjects received a minimum of some college education, their education was secular—most often in the physical sciences. They rarely, if ever, had a background in theology. He notes several extremist leaders—Qutb, Faraj, al-Zawahiri—were not religious clerics. The regression analysis in this study does not subdivide education into disciplines. Additionally, the sample size (50 events) is too small to enable quantitative analysis using subdivisions of the graduate education variable.

Crime

Crime data is represented by burglary percentage of population per department (INSEE, 2008 data). INSEE stores data on burglaries, vehicle thefts, robberies, and murders at department-level granularity. The inclusion of crime-related datasets produced autocorrelation errors in the exploratory regression model, as the data increased and decreased similarly among the datasets. The burglary dataset was selected because the regression analysis required that single representative dataset for criminal activity. Burglary numbers were normalized as percentages of the population per department. The burglary choropleth map appears to depict a positive correlation between burglaries and events.

However, the burglary coefficient map (shown below) appears to indicate an inverse correlation, as it depicts events occurring in blue areas in the region where the local R^2 values are strongest. So how can burglaries negatively correlate with events when the burglary event map appears to show a positive correlation? The answer might lie in the manner in which the coefficients temper each other's influence within the model.

To better understand the behavior of the burglary coefficient, geographically weighted regression was conducted using only burglary percentages as an independent variable. The map below depicts the influence of this coefficient on predicting safe-house locations.

Geographically Weighted Regression Coefficients Crime (Burglary Percentages by Department 2008)

Figure 11. Map Showing GWR Coefficient Strength for the Variable, “Crime” (Burglary Percentages per Population), per Department, 2008 INSEE data.

When considered in isolation, the burglary coefficient has banded regions and degrees of influence (shown in map below), and generally, this coefficient positively correlates with event locations where the local R^2 if the multi-variable model is strongest. The burglary coefficient does not exert this effect in the multi-variable model, because other coefficients come into play to adjust the weight of the burglary coefficient. The inclusion of the other variables contributes significantly to the model's accuracy, as the single coefficient GWR model has an adjusted R^2 value of only 0.174, indicating that burglaries only explain 17.4% of the likelihood of an event occurring in a predicted location.

Geographically Weighted Regression Coefficients Burglary Percentages of Population as Single GWR Coefficient

Figure 12. Map Showing GWR Coefficient Strength for the Variable, “Crime” (Burglary Percentages per Population) in a Single Variable Regression Analysis, 2008 INSEE data.

The burglary coefficient’s weak predictive ability is shown in the scatterplot below by the datapoints indicating that there are multiple departments with high burglary rates but no events.

Figure 13. Scatterplot Showing Weak Correlation between Burglaries and Events per Department

Examination of the low burglary-high raid correlation could also provide insight as to why safe-houses were discovered in high income as well as low income arrondissements in high GINI areas. Sharp income disparities within arrondissements that obscure the actual income level of a raid location might be one explanation. This explanation could neither be verified nor ruled out, because the raid reports rarely listed the raid location with sufficient granularity to make an assessment.

An alternative explanation for the occasionally inverse correlation between burglaries and raid locations concerns law enforcement's relationships with the communities they serve. This explanation posits that police have a greater and more consistent presence in high income communities, resulting in lower burglary rates. They might also have a less adversarial relationship with residents, given that burglary and robbery suspects usually live outside those communities, thus police are often conducting unpleasant interviews elsewhere. The less adversarial relationship might translate into a more informative one, where residents feel comfortable reporting suspicious activity. Therefore, police might have superior intelligence networks in higher income

communities, as residents more readily come forward to tell the police when they hear of or see something suspicious, and their information yields successful raids.

This explanation implies that there might be undiscovered safe-houses in other high burglary departments, but we are not aware of them because the residents rarely volunteer information to the police. If this explanation is correct, the fact that events are clustered in a few high burglary departments implies that we might discover more safe-houses if police had better intelligence in these areas.

Police-Community Relationships

Research on community policing and communities' relationships with local law enforcement in France does not support the hypothesis that more raids in lower crime communities is a reflection of a better community relationship generating a better intelligence network. The French government has improved local law enforcement-community relationships as a result of a 30-year effort to encourage cooperation between the two sectors, elements of 19th century policing attitudes are still pervasive. Rather than seeking to function as partners, each views the other as an outsider. The police see themselves as professionals who know their jobs better than anyone else, and thus perceive community participation in crafting and implementing policies as amateur attempts to usurp their authority. The residents tend to view law enforcement as strangers who enter their neighborhoods and demand that they come forward to report suspicious activity. The residents also resent police telling their community leaders to implement policies that are not tailored to the community's needs. French policy-makers are well-aware of the sociological, behavioral, and economic factors contributing to crime

rates and have implemented initiatives and political structures to address the problem. (Jones & Wiseman, 2006)

In 2002, the French parliament directed municipal governments to establish *Local Security and Crime Prevention Councils*, which transferred crime-fighting and crime-suppression responsibilities to the local communities. The restructuring of France's security apparatus coincided with a pilot program that provided community policing training to local law enforcement officers and interagency outreach teams with residential participation.

Unfortunately, while there has been progress, there remains much progress to be made. Police continue to design and implement policies without consulting the communities they serve, and rarely record or respond to resident complaints. Public meetings between law enforcement and residents are not well-attended, and civilians who do attend report that they hear crime statistics and directives to report suspicious activity, but are not asked for their input. Crime fighting policies continue to be designed and implemented with a top-down approach, rather than through a cooperative relationship.

However, this tense relationship is not limited to one class or community—it appears to affect white-collar, blue-collar, and immigrant communities.

Arthur A. Jones and Robin Wiseman (2006) argue that the adversarial environment creates “no-go zones” that police avoid. They assert that these are high crime areas with immigrant youth populations that feel marginalized and ignored because police won't enter to maintain order and stability. They also assert that these youth have formed

terrorist cells within these zones. However, the regression models evaluated in this research show no significant correlation with Sensitive Urban Zones (ZUS)—the so-called “no-go” zones—and North African and middle-eastern immigrant populations. Further, the regression models show no significant correlation of arrest/raid events with areas of proportionately high ethnic Muslim populations. These findings comport with those of Finney and Simpson (2006), whose research on UK terrorist suspects’ geographic origins shows that “...*Muslims charged with terrorism under UK legislation are no more likely to come from districts where the proportion of Muslims are highest, than from other districts.*” (Finney & Simpson, 2006). Many of the 752 ZUS do not have proportionally significant North African and Middle Eastern populations. Furthermore, the regression analyses revealed no significant correlation of the presence of a ZUS with an arrest or raid event. Lastly, Jones and Wiseman provide no evidence of terrorist youth gangs in these zones. If they do have quantifiable information, French law enforcement officials would likely welcome the information.

Coefficient Behavior in High Local R2 Areas

The first aspect of this analysis focuses on areas where the Local R2 shows the best model fit—the southern and then the northwestern region in Pays de la Loire. The southern areas show a weak positive correlation between burglaries and events, negative correlations between income and events, but a mix of positive and negative correlation with respect to the correlation between graduate degree holders and events. This behavior counters much of the conventional wisdom in security studies theory. Keeping in mind that correlation does not necessarily mean causation, insights into these coefficients might be gleaned from an examination of the events recorded in these

areas. Every incident with the exception of one event in Albi occurred in Toulouse. Several arrests were connected with the investigation into Mohammad Merah, and 17 arrests were connected to suspected followers of Forsane Alizza chief Mohamed Achamlane.²³

In Toulouse, income and education coefficients exert a weak positive influence on event probability, but crime exerts a weak negative influence. According to the model, the more education and income a Toulouse resident has, the higher the probability the resident will tolerate or support a haven for extremists. The choropleth map of average household income by per department highlights the economic contrast of Toulouse with the surrounding departments.

²³Bruguière provides details on the various networks active at the time in Jean-Louis Bruguière and Jean-Marie Pontaut, *Ce Que Je N'ai Pas Pu Dire*, (Paris: Laffont, 2009), pp.443-463.

Average Declared Departmental Household Income 2008*

*Toulouse highlighted in Turquoise

Figure 14. Average Declared Departmental Household Income, INSEE 2008 Data.

However, it would be premature to make the analytic leap that in Toulouse, the wealthier a resident is, the more likely he is to look out the window over an evening glass of merlot and observe his neighbors unloading strange cargo. The news articles reporting on terrorism-associated raids in Toulouse did not provide detailed location information, so at the time of this research, I was unable to determine exactly where in Toulouse these raids occurred. The events could have occurred in low-income areas because Toulouse has a high GINI index. The wealthier residents of that department drive up the income level, but there are also areas that are much less well-off, sometimes in close proximity to upper income neighborhoods.

Toulouse contains gated communities that exist side by side with low-income neighborhoods. Residents on opposite ends of the income spectrum often live in close proximity, yet are highly segregated. They shop in different stores and do not typically cross paths. The department is has also grown very rapidly over the past 10 years. Although job opportunities for business executives have increased, blue collar jobs and independent contractor opportunities in Toulouse are shrinking.

The social segregation and gating of communities is driven by security concerns on the part of those living in the upper income areas. Robert Monné, CEO of Monné-Ducroix, a major developer of gated communities in Toulouse, explains that while people value security, no city has the resources to provide complete security, gated communities can bring residents peace of mind. The transition to gated communities follows the American model. For Americans, this choice is consistent with a nation that is largely capitalist and unabashedly multi-class. In France, it can seem a bit incongruous and hard to justify for more socialist-minded populations. As new residents of one of the Toulouse gated communities explained to a journalist covering the trend, 'Even if we still vote for the Communist party and are against the idea of social segregation, we want to live somewhere quiet'. This is not unusual behavior. Monne adds. "...in fact, most of the councils where we have located estates are on the left..." (Belmssous, 2002)

Not surprisingly, the positive income coefficient coincides with a positive graduate education coefficient, since high education often provides the skills that command higher salaries. Unfortunately, without more specific data on where the raids occurred,

interpretation of the coefficients' impact must be tempered to more general analyses, as background information on those arrested in connection with the event data points is difficult to uncover.

VI. Background of suspected extremists—Do They Fit the Model?

Some of those arrested do have criminal backgrounds. For example, according to French officials, Muhammed Merah, the French national who in 2012 committed several terrorism-related killings in Toulouse and Montauban, came from a working class neighborhood and had a history of petty crime arrests. Merah was associated with the Forsane Alizza network and trained at terrorist camps in Pakistan and Waziristan. Merah does not appear to be well-educated or come from a high-income family, though he does have a petty criminal history that would be consistent with the generally positive crime rate coefficient.

Another Toulouse-based network was associated with Mohamed Merah, who conducted terrorist attacks in March 2012 that resulted in seven deaths, four at a Jewish school in Toulouse. This group conducted attacks in southern France, and some group members also fought in Syria. (Holman, 2015) Police charged two group members as Merah's accomplices. One individual was, according to police statement, "a member of the traveler community who converted to Islam..." The other group member charged with assisting Merah was the traveler's ex-girlfriend. (Phelan, 2012)

Information about the traveler's educational background and income level was unavailable at time of research, but travelers tend to have limited income and education in France as well as in Bulgaria and Romania, countries from which many of them emigrate. (Chiriac, 2012) Friction between travelers, also called Roma, and government authorities has sometimes become violent. In July 2010, French Roma attacked a police station, destroyed trees and set cars on fire in the Loire Valley in retaliation for the police shooting of a Roma who knocked down a police officer as he drove through a police checkpoint in the process. President Sarkozy responded by ordering the dismantling of approximately 300 illegal camps within three months. He stated that the camps were "...sources of illegal trafficking, of profoundly shocking living standards, of exploitation of children for begging, of prostitution and crime". (BBC, 2010) In 2009, authorities implemented a mass deportation of Romanian and Bulgarian illegal immigrants, expelling them from camping sites throughout the country. (BBC, 2010)

One of the people arrested in a March 2012 raid does not appear to hold an advanced degree or to be financially comfortable. Willie Bridgette spent much of his time sharing boarding with others, some of whom shared his extremist views. He is now serving a nine-year sentence. Forsane Alizza member Willie Brigitte lived in Couronnes, a suburb in the 11th arrondissement of Paris, converted to Islam and attended mosques and Islamic schools with reputations for preaching extremist anti-Western messages. He twice deserted the French navy and held a variety of jobs including butcher, social worker teacher, and drug rehab counselor. He was twice divorced and had three children. Police suspect him of obtaining travel documents for Ahmed Shah Massoud's

assassins. In 2007, he was convicted of planning to attack an Australian nuclear facility. (Peterson & Fray, 2004)

Some of the individuals linked to the event data were well-educated and possibly financially comfortable. Rany Arnaud was arrested in December 2008 for planning to plant a bomb outside a government building near the Eiffel Tower. Arnaud was educated in computer electronics and lived in Val-de-Marne, a moderately well-to-do area just outside of Paris. Rany traveled to Syria and made several attempts to join the insurgency in Iraq. He eventually joined Algerian radicals. Upon his arrest, he informed police that he planned multiple attacks on French soil, including the French Intelligence Service headquarters. Arnaud was arrested with Nadir Badache and Adrien Guihal, two suspected co-conspirators. I was unable to find information on the backgrounds of Badache or Guihal.

I could find relevant background information for one of the Paris cases—that of Jeremie Sidney—though the information was insufficient to provide insight on whether or not the case conformed to all three model coefficients. In October 2012, Jeremie Sidney and another man threw a grenade at a kosher supermarket in Sarcelles, a Paris suburb. Jeremie Sidney was radicalized while in prison for drug trafficking. He joined a jihadist group and recruited fighters for the insurgency in Syria. (Hinnant, 2012) Jeremie Sydney did not appear to be well-educated or well-to-do, but his case does conform to the model in that he has a criminal history.

VII. Factors excluded from the regression analysis

Social Environment—Impact of Ethnicity

In order to assess the impact of factors associated with supportive populations and historically marginalized social environments, it is necessary to identify areas with high North African and Middle Eastern immigrant populations. Data on ethnic population percentages were included in the exploratory regression analysis to assess their correlation with event locations. A high positive correlation would indicate that these ethnic communities might be more tolerant of extremist ideas than are areas that have other ethnic compositions or predominantly of multi-generational French heritage. Extensive searches for GIS and tabular datasets detailing populations of individual ethnic groups, immigrant nationalities, and religious populations were not fruitful. INSEE maintains datasets on immigrants vs. non-immigrants and French citizens vs. non-citizens, because France imposes legal restraints on collecting this type of data.

The French government must conform to Article 8, Law number 78-17 (6 January 1978), a law that prevents the collection of data on ethnicity and religion on the grounds that the collection of such data is an invasion of privacy. The spirit of the law is traced to regulations enacted in 1872 that prohibit the government from collecting personal information on ethnicity and religious affiliation. (Camilleri, 2012) The most specific datasets I was able to find address ethnicity on the department level, giving population numbers for first-generation Algerian, Moroccan, Tunisian, and Turkish immigrants in separate fields. I summed these fields and computed the percentage with respect to the

total population for that year per department. The result is, hopefully, a reasonably accurate representation of the proportion of northern African and Turkish immigrants per department.

I also compiled a table that lists the number of mosques per commune, including the address and any descriptive notes on the size of the mosque and the theological bent of its imam(s). For example, if the imam advocated progressive ideas or strict, unyielding interpretation of Islamic laws. I abandoned this compilation because it became impossible to determine the theology of many of the mosques. Additionally, it became clear that an accurate representation of the concentration of mosques would be daunting, and its accuracy impossible to verify.

Societal Membership—and the Political Voice

Support of the head scarf ban was much stronger among the broader French population (69%), and may indicate a significant risk factor that is much harder to quantify in a model—societal membership. Researchers from Sorbonne University and University of California, San Diego, and Stanford University found that job applicants with an ethnically Muslim name were 2.5 times less likely to receive calls to interview for a position than were candidates with comparable CVs but non-ethnically Muslim names. Researchers controlled for regional differences, employment sector, company size, occupation, contract type, and whether or not the CV included a photograph.

If societal membership is an inverse risk factor, then how might we quantify it for a model? One metric might be Muslim representation in local politics. The table compiled for this research records the proportion of council members of possible North African or Middle Eastern ethnicity in the commune they represent. Representation is difficult to quantify, as representatives are not required to identify themselves as Muslim. Therefore, the numbers are determined based on the council member's name.

The relationship between the percentage of local council representatives of North African or Middle Eastern ethnicity and the number of arrests/raids in those areas between 2008 and 2012 is expressed in an R^2 value of .0295. The graph below shows the relationship between events and ethnic representation on the municipal level. There appears to be no linear relationship, but the graph will be regenerated after the datasets are verified for accuracy.

Figure 15. Number of Events and Percentage Middle-Eastern and North African Municipal Representation in 2008.

Outlier Communes

Aubervilliers is an outlier in the data, in that the commune has the highest ethnic representation of any of the communes in the dataset, yet there was one event within its boundaries. However, removing it from the dataset did not have much impact on the correlation. Recalculation after removing the outlier still returns a weak correlation. Thus, based on these metrics, there is a very weak correlation between social representation, at least in the political arena, and the presence of safe-houses.

Proximity to Sensitive Urban Zones

While not all 752 sensitive urban zones had arrests/raid events, most recorded events occurred in or within a mile of communes with sensitive urban zones.²⁴ Initial research identified 512 communes containing at least one ZUS.²⁵

²⁴The average area of a commune is approximately 5 square miles.

²⁵ Number is approximate due to differing definitions of what area is a commune.

Sensitive Urban Zones and Arrest/Raid Locations

Figure 16. Map Showing ZUS Regions and Event Locations.

A ZUS is an official designation for what French authorities deem an “at risk” area. According to INSEE, France’s National Institute of Statistics and Economic Studies, Sensitive Urban Zones are “Les zones urbaines sensibles (ZUS) sont des territoires infra-urbains définis par les pouvoirs publics pour être la cible prioritaire de la politique de la ville, en fonction des considérations locales liées aux difficultés que connaissent les habitants de ces territoires.” “Sensitive Urban Zones (ZUS) are sub-urban areas defined by the government to be the primary target of the policy of the city, according to local analyses of the problems known to inhabitants of these territories.” ZUS are

characterized by high unemployment and low high school graduation rates. There are 752 ZUSs in France, 718 of which are on the mainland in metropolitan France. There are 36,552 communes in metropolitan France, and 490 of these communes contain at least one ZUS. (Leonard, 2015) (Henry & Dieusaert, 2014)

In 1996, the French government created the Sensitive Urban Zones as part of their Urban Regeneration Pact. The Pact sought to bring public services to urban communities, ensure the security of residents, stimulate the local economy to increase employment opportunities, and promote social diversity. The government established an Economic Revitalization Fund to financially support this program. It initially designated 700 ZUS and added other zone types—350 Urban Regeneration Zones (ZRU) and 38 Urban Free Zones (ZFU).

The Urban Regeneration Pact increased police presence, and special educational facilities were established to accommodate juvenile repeat offenders.

In 2003, the French government augmented ZUS funding when it established an Urban Renewal Programme for ZUS. This program had a five-year, €30 billion budget earmarked for the most blighted neighborhoods, and would be applied to physical improvements and attracting businesses to the areas. The program lures businesses to these areas with special incentives, such as waiving social security contribution requirements. A portion of the budget goes to special skills training for residents in need of education to make them more employable. (Government of France, 2007)

Communes that contain at least one ZUS:

10 of the 50 events could not be identified at the commune level. For the 38 events that were pinpointed to a commune, 36 of those occurred in communes that contained at least one ZUS or were themselves a ZUS. The actual percentage may be as high as 94%, given that 9 of the arrests/raid locations could not be located with more specificity than the general locale of Paris, and one (a multi-arrest/raid in Herault) could not be identified at the commune level. If this is the case, then the presence of a sensitive urban zone is a strong predictor of safe-house location.²⁶

The correlation between the number of proximate ZUS and event locations increased when ethnic population was incorporated into the linear regression analysis. Linear regression conducted on ZUS against arrests/raids showed a Pearson's correlation of .777.

If the number of proximate ZUS was considered with the percentage population of ethnic origin (Algerians, Moroccans, Tunisians, Turks) per department, the Pearson correlation index increased to .8, indicating a high degree of correlation between these variables and the number of arrests/raids. Therefore, preliminary analysis indicates that ethnic composition does not significantly increase correlation, but the proximity to a ZUS is a key indicator of an area's risk.

²⁶ The news articles that report of arrests/raids in Paris do not specify the arrondissement in which the arrest/raid events occurred. To include the events and analyze commune-level data, Paris is treated as a single commune for the purposes of this analysis.

ZUS Components

Given that the presence of ZUS is correlated so highly with arrest/raid locations, the question logically follows, “what is it about a ZUS that makes it correlate so highly with the occurrence of arrest/raids?” Since the French government uses high unemployment, high crime, and low educational attainment as the main criteria determining designation, it makes sense to begin by examining these phenomena individually for any correlation with arrests/raid events.

Unemployment

According to Exploratory OLS and GWR applied to the France test case, unemployment is a very weak predictor of events, scoring at 3.12% significance in the exploratory model, just above the weakest predictor tested, the percentage of residents between 15 and 64 years of age with no education beyond primary school. This assessment appears to counter a body of social science analysis that argues otherwise. An IRIN News analysis cites unemployment as one of three significant radicalization factors in Kenya: “Unemployment, poverty and political marginalization are contributing to the Islamic radicalization of Kenya's youth, ...” (Integrated Regional Information Networks, 2013) There are a number of potential explanations for its lack of predictive power in this model. High unemployment might be an influential factor in Kenya, but not in European countries, or it might only be impactful only if accompanied by other risk factors. Aggravating factors might be a high unemployment rate combined with high youth percentage in population. High GINI index might be another. The regression analysis conducted in this research, however, does not indicate that either factor combined with

unemployment constitutes a combination that encourages radicalization or an environment vulnerable to the establishment of extremist safe-houses—at least in the case of France.

Secondly, the unemployment variable might only exert an impact if the unemployment rate is very high. The point at which unemployment becomes an aggravating factor is a separate empirical research topic in itself.

Average Unemployment by Department in 2008

Figure 17. Average Unemployment by Department, 2008 INSEE data.

Low Educational Level

The correlation between radicalization and limited education has largely been dispelled by recent research. The assumption that low education is linked to radicalization persists in professional circles, despite over 30 years of counter-evidence. Russell and Miller conducted research in 1977 to develop a sociological profile for terrorists. Based on statistics for over 350 terrorists in America, Europe, Asia, and the Middle East, they built a composite profile of a typical terrorist as a single male in his early to mid-20's with a university education. Interestingly, most of the terrorists in the study were recruited while pursuing their university education. A notable difference between Russell and Miller's discoveries and that of more recent research is that the majority of extremists in the 1977 study had academic backgrounds in the humanities, law, and medicine, whereas a study conducted in 2009 indicates a majority pursued careers in physical sciences, such as engineering. Russell and Miller detected this shift and recommended further research to refine the profile to reflect "...increased involvement by technically trained individuals,..." (Russell & Miller, 1977) Gambetta and Hertog's 2005 study found that university graduates with physical science backgrounds were "...strongly overrepresented..." in Islamic extremist groups. Gambetta and Hertog compiled educational data from 404 perpetrators of politically-motivated violent acts in 30 countries between 2005 and 2007. Of the 178 perpetrators whose academic backgrounds could be determined, 44% had majored in engineering. Further, no other academic specialization followed closely behind the engineer majority. The second-largest group consisted of Islamic Studies majors, who comprised 19% of the sample set. Gambetta and Hertog do not offer an explanation for this high proportionate representation beyond their observation that engineers tend to hold more conservative

views. Even if this observation is accurate, it does not adequately explain the tendency toward radicalization, as extremism can be interpreted as either highly conservative or highly rebellious behavior, depending on context or perspective. However, during an interview by IEEE Spectrum magazine, Hertog observed that “Engineering seems to attract a larger share of people drawn to rule-bound systems, compared with other scientists who primarily work on open-ended questions and might be more skeptical.” (Hertog, 2008) In fact, empirical research indicates the opposite might be true. As Berrebi concludes from his study, “If anything, the correlation I find is that those with higher education and higher living standards are more likely to participate in terrorist activity.” (Berrebi, 2007) As explained earlier, Hassan’s (2001) study of 250 Palestinian extremists revealed that none were uneducated.

The variable expressing limited or absence of formal education was included to investigate whether or not France exhibits outlier behavior. In short, it comports with the contemporary analysis in that events occurred in locations and involved individuals with limited as well as high education.

Exploratory regression analysis that included this variable found the low education coefficient statistically significant for only 27% of the datapoints. Its influence was evenly split, with 55.21% of the arrondissements exhibiting a positive correlation between limited education and events, and 44.79% of the arrondissements showing a negative correlation.

Percentage with No Education Beyond Primary School by Department in 2008

Figure 18. Population Percentage Aged 18-64 with No Education Beyond High School by Department, 2008 INSEE data.

State Policy

For the purposes of this research, state policy was an initial candidate variable because states can implement policies that affect a location's suitability for harboring violent extremists, either intentionally or unintentionally. For example, they may do so intentionally through the use of terrorist entities as foreign policy tools, perhaps providing training and boarding locations to the groups that advance the host state's interest.

States can also be unintentional sponsors if they are controlled by a weak central government, or if they tolerate—or cannot control—government corruption at the local level and radicalized elements take advantage of that vulnerability.

Some states—Syria and Iran, for example—have been accused of using terrorist groups to achieve foreign policy ends. Before the 2011 conflict in Syria, members of the international community asserted that Syria sought to further its interests in Lebanon by directly supporting Hamas and by working with Hizballah. Hizballah, in turn, has been linked to Iran, with whom it shares a friendly relationship and religious ideology. Such ties could extend to offering those groups safe harbor and training. Such was the case of Pakistan's early assistance to the Taliban, which is now threatening its stability both directly and by complicating many of its diplomatic initiatives. Intentional state sponsors are safe-houses on the macro-level, and groups that have the support of governments can strike far beyond neighboring states and be truly global, with considerable resources from which to draw.

Unintentionally hospitable state policy can have significant impact as a variable coefficient simply because terrorists are more likely to establish bases in regions within which it is easiest to operate. This phenomenon accounts for much of the resiliency of terrorism. It expresses on a macro-level what law enforcement faces on a micro-level whenever they attempt to eliminate drug-dealing and prostitution from an afflicted neighborhood—the crime is not eliminated, as offenders simply move to a more hospitable region and continue their activities.

State Policy—Identity Conflict Dynamic

This political environment can exist in countries containing multicultural regions that are often near a border, where residents may identify with the neighboring state rather than the state in which they reside, thus compromising the allegiance of local officials who are often residents of that enclave, as well. The importance of social identity over national identity can diminish the respect of legal boundaries, rendering certain illegal activities socially and morally acceptable. Some Muslim enclaves in France emphasize their social identity over that of their French citizenship. Radicalized French natives identify with those social groups whose extremist doctrine they adopt. Thus, they not only identify with that social group, but with the larger culture or nation from which it originated.

In some instances, social identity not only isolates the group from the larger community, but also shapes its attitude toward the government. Local residents do not view the host country as their supporter or protector, and they often view the state as “other”, rather than as a macro-expression of their local society. This attitude can be found in many Muslim communities in France. These residents do not identify with native French citizens and by extension, the French government. They believe they are viewed as outsiders, and that the government holds the interests of the native French over their own. Their perception is reinforced by domestic policies that restrict Muslims, but have no similar effect on other groups. Thus, their attitude of “otherness” toward the state moves beyond characterization as merely foreign and becomes adversarial.

The central government's influence is further weakened if the region is also geographically remote. This is often the case in regions of Afghanistan, Pakistan, and in Syria before its current internal conflict. This dynamic does not apply to France, but it is an important variable to consider when applying the model to other cases.

A state's political environment shapes many other variables, generating domestic and foreign policies that affect its cultural environments and its physical security (e.g., enabling or inhibiting smuggling routes). For example, France's refusal to participate in the overthrow of Saddam Hussein and its active support of Palestinian statehood demonstrate that it is not generally hostile to Muslim views, or even to some positions championed by religious extremists. Neither is it an autocratic state that violently crushes all dissent. However, French perspectives on identity create difficulties in reconciling one's identity as French for those whose cultural and religious identities originate from a region beyond French borders. Ascertaining the relative importance of these two factors—the weighting of variables in multi-modal modeling systems—is critical to the determination of a tipping point with respect to influence of identity on a location's vulnerability to hosting extremist safehavens. Successful factor weighting can also determine whether or not changes in public policy or foreign policy can sufficiently alter the weighting and improve the security situation.

State Policy—The Headscarf Ban

State policy could be a contributing factor because some policies impact daily life and can even send an official message about what the state values in terms of the social identity of its residents. France's 2005 prohibition of head scarves in public schools is

one such example. The ban implemented Laïcité—the French policy of secularism. The government argued that wearing a headscarf in school amounted to displaying a religious symbol in a public setting. Most French citizens supported the headscarf ban, and Muslims might interpret this widespread support of an official prohibition of a religious practice as a rejection of their religious identity. The message in effect being, “you aren’t part of the larger ‘we’, and we don’t want to see you in our public spaces”. However, Laïcité as a state policy might not be a risk factor. More likely, the actual risk factor is *the popular attitude that supports the policy*. In 2004, Comprendre Pour Mieux Décider (CSA) conducted a survey to gauge the level of support for the ban before it was implemented and found that 49% of Muslim women supported the ban. A significant proportion (43%) opposed it, but this split suggests that if the headscarf ban was intended as a societal message, it was not a clear one. (Evry, 2004)

Supportive Population

A supportive population is necessary for the group’s sustained functioning as a macro-safe-house. The ethnic population also provides some of the group’s funding and is usually the primary recruitment base. Thus, the group’s ability to identify with the ethnic population is a prerequisite for its continued operation. Successful identification does not rest solely on shared culture. Agreement with terrorists’ grievances must be shared by the ethnic community, and that community must believe that terrorist group advances their shared interests.

The expression of this contributing factor becomes more complicated with the existence of multiple groups whose interests overlap and who both claim the same cultural identity. Although it is possible for more than one terrorist group to enjoy the support of an ethnic community, resources are a valuable benefit of securing popular support. Money and recruits are finite resources, and competition between terrorist groups often occurs. The groups compete for the community's devotion by engaging in activities designed to demonstrate their dedication to the shared cause and their value to the community. Dedication is sometimes demonstrated through launching operations designed to draw attention on the world stage. However, such endeavors require some care, as it is possible to overstep the bounds of what is acceptable to the population. Value to the community is expressed by providing services that the population appreciates. Hamas and Hizballah have conducted many terrorist acts, but also expend resources within their ethnic communities to provide social services in order to demonstrate their value as community benefactors (McCauley & Moskalenko, 2008).

Religiosity Coefficients

This research seeks insight into the locations that radicalized individuals consider suitable to establish residence, recruit, store supplies, and plan attacks. The individuals associated with the test data claimed religious motivation based on radical misinterpretation of Islamic teachings. While variables quantifying supportive populations (detailed in the previous section, "Supportive Population") sometimes touch on religion, another variable was designed to quantify radicalized religiosity. This "radical religiosity" variable—RR coefficient was designed as a compound variable that incorporated the

concentration of mosques with the mosques' interpretations of Islam. The process involved researching religious leaders who espoused intolerant, extremist, and finally, explicitly violent interpretations of Islamic concepts. Once identified, further research was conducted to identify the mosques they led and where they were welcomed as guest clerics. Another variable for "inclusive religiosity"—IR coefficient—was developed for locations with mosques led by clerics who espoused an unusually inclusive interpretation of Islam. For example, Mohammed Ludovic Lütfti Zahed encourages tolerance of homosexuality, a behavior that is generally rejected by Muslims. (Würger, 2014) The IR variable could be negative or positive coefficient. It might lessen the attractiveness of a site if a radicalized individual perceived the area as an unlikely area for successful recruiting, or as an area where the risk of being exposed to counter-terrorism authorities was high. Conversely, it might function as a positive coefficient if the Imam were in the minority, and the surrounding population was deeply offended by the unusually liberal interpretation of Islam.

To build the shape file for this variable, a dataset consisting of the names of Mosques, their geographic coordinates, addresses, and associated leaders was developed from extensive internet research of newspaper articles, blogs, interviews, and academic papers that addressed Islamic practices in France. The variables were judged inapplicable to France for several reasons. Firstly, quantifying the mosques was daunting based on sheer volume. As of 2009, there were over 2,000 mosques in mainland France. (Marcel, 2009) It is difficult to obtain an exact number, largely because the fundamental question of what constitutes a mosque is difficult to answer. This question could not be answered satisfactorily given the data available, even when a clear definition is established. While conducting research for each mosque included in

the dataset, it was discovered that some entities listed as mosques were little more than small prayer rooms. The question arises, “isn’t that a mosque, as well?” One could define a mosque as a place of prayer guided by an individual acknowledged in religious circles as qualified to lead prayers. If so, then many of these prayer rooms are mosques. The legitimacy of these sites as mosques could not be established if the identities and credentials of their imams were not explicitly available, and this was often the case. Even many of the larger mosques did not list the identities of their religious directors. Additionally, even if this data were available, its inclusion might introduce a temporal component in the analysis. The imam’s tenure at the mosque would have to be temporally established and bracketed as a variable, and the residency period of the sample set subjects would have to be established and incorporated, as well. Temporal analysis would be a welcome addition if the sample size were large enough to incorporate temporal analysis into the regression model. Since there were only 49 events from 2008-2012, enlarging the sample would require either expanding the time window, or engaging French counter-terrorism and law enforcement authorities to enhance the sample set granularity.

Existing Smuggling Routes

Transportation routes for the support of terrorist activities are often established along pre-existing routes used for other illegal activities, such as drug-smuggling, and terrorist operations are frequently funded through illicit trade. For instance, France is a major transit and distribution point for Moroccan cannabis, thus there is a pre-established network for transportation related to terrorist activity. (Caron, 2015) (Boekhout van

Solinge, 1995) Therefore, it is not surprising that convicted terrorist Menad Benchellali had a brother Hafed who was convicted of a financial crime—payroll theft. Their father's connection with the Algerian Armed Islamic Group (GIA) has provided a much needed avenue through which to obtain financing. The group used its ties to extremist enclaves in France (and elsewhere in Europe) to engage in drug and weapons trafficking. Thus, not only does existing illicit trade increase the probability of the development of terrorist safe-houses, but the presence of these locations increases the local crime rate, as the town of Roubaix experienced when its residents fell victim to several armed robberies.

A 2005 report by Shelley, et al. (2005) entitled *Methods and Motives: Exploring Links Between Transnational Organized Crime & International Terrorism*, studied links between criminal networks and terrorist groups and identified 12 situations, or watch points, in which this interaction was likely to occur. Such instances result in the enhancement of both criminal and terrorist entities. The coincidental needs of both sectors (i.e., safe-houses, smuggling routes, untraceable funds) highlight similarities between the two groups, thus some contributing factors are held in common. These factors closely resemble the watch points described in the report.

Although the research of Shelley et al. primarily focused on criminal organizations, the ability of these groups to enhance terrorist operations warrants that the significance of social environment to criminal networks be briefly addressed. Criminal networks often identify with their parent culture and use this affinity to deeply entrench themselves within their ethnic neighborhoods, and they do so on several levels. Firstly, they apply their intimate knowledge of the social mores, customs, and language of the culture to

garner sympathy from their community. Secondly, they capitalize on the insular behavior of an ethnic community that is disenfranchised from the mainstream population. The community either supports them as “one of their own”, or fails to report their crimes out of mistrust of authorities and a lack of confidence in its ability to protect them in what is, in some ways, a state within a state.

Border Proximity

Geographical proximity to inter-state borders and transportation routes increases the vulnerability of the locale. These locations allow actors to establish stable bases without sacrificing the mobility required to commit inter-state acts or seek refuge if threatened by local authorities. Proximity to borders and transport routes (both inter-state and local) facilitates the movement of illicit goods, such as weapons or drugs.

Venissieux, France is an example of a city offering a supportive environment near a state border, providing easy access to major highways and less than two hours from Switzerland. Local residents of this Lyon suburb identified more closely with radicalized individuals in their midst and actively supported them by falsely reporting their passports as “lost” so that a local extremist cell could use them for interstate travel. Venissieux residents Mourad and Nizar Sassi used the documents to enter Afghanistan to attend an Al-Q’aeda training camp, where they were captured while combating U.S. forces. (Associated Press, 2004)

Roubaix is another example of a French locale with a supportive population that is proximate to an interstate border. It is just over an hour from Brussels. In 2005, Roubaix resident Lionel Dumont was sentenced to 30 years in prison for his role leading an Islamic extremist cell that financed jihadist attacks via armed robberies. (Mili, 2006)

Border proximity was not included in the exploratory regression analysis because France's extensive road network, easy access to airports and sea ports render interstate travel convenient for virtually every French locale. It is however, worthy of note that 14 of the 50, or almost a third of the events in the dataset are positioned within eight miles of an inter-state border or sea coast. This might seem like an unimpressive proportion, but it implies that geographic proximity to a border could have impact in other case studies.

Micro-level Variables

One of the most frustrating aspects of using quantitative analytic techniques is that the analysis is often limited by the availability and format of the data required for the analysis. Micro-level variables were omitted from the exploratory regression analysis because the data granularity simply was not available in this research context. While the demographic and municipal data exist at the required granularity, the test data are not openly available to this level of specificity. Thus, while the GIS files can be constructed—no easy feat, but not impossible—the test data could rarely be resolved to the specific address. Since there are only 50 datapoints in the test dataset, nearly all of them would have to have building level location data as a prerequisite to including micro-level

variables in the model. Greater refinement of the test data could result in a more nuanced model that incorporates these variables.

While the previously described factors can influence neighborhoods at more local scales, additional factors can manifest on the micro-scale that impact the selection of the base or safe-house itself, once a local neighborhood is identified.

Desirable structures exhibit several characteristics. The ideal site is surrounded by space sufficient for maneuverability while hidden from view. For instance, the presence of buildings with high surrounding walls, trees or other objects that obstruct aerial view, enable the transport of contraband and operatives while unseen. This area provides a protected space for the alteration of vehicles for purposes such as obscuring the vehicle's identity from law enforcement or the installation of weapons.

Extensions of the geographical aspect of this factor include the presence of hidden driveways. Angles or structures that enable one-way surveillance, such that the group can be aware of others approaching without themselves being seen are also desirable characteristics of safe-houses. Therefore, buildings that exhibit these features should be viewed with suspicion as buildings likely to serve as safe-houses.

Access routes are another important consideration and their viability depends on geographical and social factors. Proximity to routes that lead to borders facilitate either cross-border attacks to neighboring states hostile to terrorists or an escape to states out of reach from local law enforcement. Major thoroughfares that lead to mass modes of transportation, such as airports and train stations, enable increased mobility, sometimes

with the possibility of anonymity, as some forms of mass transit, such as busses, trains, and busy streets, allow individuals to evade law enforcement and become “lost in the crowd”.

Social factors that contribute to the desirability of access routes include supportive local populations, populations of diverse ethnicity, and high-crime areas. The importance of supportive populations is described in the section specifically addressing this factor, but it is important to note that although a supportive population exerts an influence at both macro- and micro-levels, its impact is usually felt most acutely at the micro-level. The effectiveness of a supportive population stems from its “grass roots” power. If the neighborhood sympathizes with the terrorist group or is at minimum, hostile to government officials, they will be aware of the movements of the individuals, but allow them to continue operating by refusing to provide information leading to the group’s apprehension. In extreme cases, such support effectively extends the safe operating space if the terrorists can be seen on the streets with little or no security risk to themselves.

As is often the case for research involving the social sciences, efforts to establish precise and discrete objects of study prove challenging. More importantly, when achieved, such divisions distort results, as they omit qualities that can only be observed if the social mechanisms are viewed holistically. The interdependence among many, if not all, of the contributing factors is readily apparent. Yet, their ability to vary as separately identifiable elements is also clear. The ability to perceive the subtle relationships and the changes that their variations exert on the whole is critical to producing a meaningful model with practical applications.

VIII. Conclusion

This dissertation research demonstrates that if spatial components are omitted, the independent variables have a weak to insignificant correlation with the presence of safe-houses and often, multicollinearity is present when multiple variables are considered in combination. However, the incorporation of geographic weighting reveals linkages between independent variables and the presence of safe-houses because geographically weighted regression integrates spatial characteristics that affect social phenomena. A cornerstone of GIS theory is Tobler's First Law of Geography which states that, "everything is related to everything else, but near things are more related than distant things". Geographic regression analysis incorporates the social reality that those of similar social characteristics tend to cluster together, and as the distance increases from a point of comparison, those similarities tend to diminish. An analytic technique that can account for this phenomenon has greater analytical power than a technique that only graphs quantifiable relationships when applied to social topics.

The factors presented for this model are widely recognized by specialists in security studies as playing an influential role in terrorist activity. Though this research has provided insight concerning these factors, debate on their relative importance, degree of interaction and interdependence is expected to continue, as it should, especially considering the data challenges associated with this case study. Future efforts building upon this research can refine the model design and apply it to more extensive datasets to provide a deeper understanding of these issues, and this information can be useful to counteract radicalization and identify potential locations for safehavens.

Furthermore, the model's GIS environment facilitates comprehension of the results by providing a visual depiction of the output, potentially presenting the user with multiple levels of analysis within a single graphic. Thus, it can convey results in a manner unmatched by a strictly textual format. This feature is particularly useful when conveying analysis to those lacking a background in quantitative analysis.

This prototype model can serve as the foundation to develop a model that not only informs security specialists on a practical level, but also offers theoretical and policy-focused benefits. Geographically weighted regression is flexible and scalable in that it can incorporate multiple variables and process large datasets, producing a customized design for the region to which it is applied.

With the integration of additional relevant datasets of sufficient granularity, the model could provide indicators of at-risk areas. If test data with high demographic granularity is included, micro-level variables excluded from this prototype (described in chapter "Micro-level Variables") could be evaluated to characterize their impact. For example, specific addresses of those arrested and of raid locations could test neighborhood and building characteristics with the inclusion of datasets with similar specificity. INSEE stores this type of data.²⁷ Government agencies and non-profit organizations often have specialized databases that could be incorporated in the model to evaluate other variables.

²⁷ An example of this type of high granularity dataset is base-ic-logement-2011.xls. This file contains the number of rooms and type of residence, how many reserved parking spaces, type of heating, the year it was built, how long residences were occupied before the residents moved, and other variables, all on the arrondissement level.

Refinements to this model can assist policy-makers and social scientists, providing insights into whether state policy on the local level is more influential than at the national level, or whether those policies can provide an effective counterbalance to other factors, such as immigration waves, demand fluctuation of employment skillsets, and under what conditions those policies would have the most impact.

For example, in 2008, France's Ministry for Higher Education and Research launched Operation Campus to improve France's international standing in the higher education and scientific research arenas. The plan allocated €5 billion to selected universities, enabling the establishment of 10 "super-campuses" for higher education and research. University recipients were chosen on the basis of "...their pedagogical and scientific ambition, building renovation needs, proposals for improving student life and the extent to which their region would benefit from their 'structural and innovative character'". (Marshall, 2008) A portion of the funding established "super-campuses" in Lyon and Toulouse, areas that showed positive correlation of higher education with terrorism-associated raids or arrests in the prototype model.²⁸ Higher education showed no significant correlation in Strasbourg, also a super-campus recipient. If a refined version of the model with higher data granularity showed similar results, the funding might be more effectively allocated to a region where higher education was a negatively-correlated coefficient.

²⁸ To prevent misinterpretation of this paragraph, I re-iterate that the model developed in this research is a prototype. This paragraph does not imply that 1) there is a definite correlation between higher education and the existence of a location that hosts extremists, or 2) that French policy-makers should have made different policy decisions. First, the model-due to limitations mentioned in the body of this text-does not include datasets of sufficient granularity and scope to assess with strong confidence that the three coefficients identified in the model are in fact, the three most significant variables, nor that the nature of their correlation-negative or positive-is reflective of reality.

The geographically-weighted regression included in ESRI software is based on ordinary least squares statistical method. As the research developed, datasets that would be the basis for coefficients grew very large, especially in relation to the test data (events). The problem of the small sample set was compounded by granularity issues that became more pronounced as the research progressed. These problems were highlighted through extensive testing using different types of regression models (including logistical regression). While GWR provided insight into the research question, its use of ordinary least squares to generate the weights proved to be a less than ideal fit.

However, other regression techniques can be applied in future research to yield more accurate results if these techniques can be integrated with geographic weighting. For example, the application of regularized regression techniques would address problems associated with overfitting, a vulnerability associated with this dissertation research, given the volume and complexity of the independent variable set relative to the sample size.

Lasso regularization constrains the residual sum of squares to the parameter vectors' L1 norm to a maximum value. The L1 norm minimizes the sum of the absolute differences between the observed and estimated values. Robert Tibshirani developed this regularization technique and uses a Newton-Raphson algorithm to perform an iterative reweighting least squares loop to determine the maximum value. (Tibshirani, 1996)

The Newton-Raphson algorithm employs a function based on the maximum value from a Taylor series approximation. The Taylor series approximates the function that is the

closest fit to the test data's behavior against a set of parameters. The resulting maximum constant is used to specify the algorithm for first-order optimization.

For the Lasso algorithm, if x is the independent variable and y is the dependent variable, the lasso estimate (α', β') is

$$= \arg \min \left\{ \sum_{i=1}^N (y_i - \alpha - \sum_j \beta_j x_{ij})^2 \right\} \quad \text{subject to } \sum_j |\beta_j| \leq t.^{29}$$

where β is the parameter coefficient. The sum of the absolute values of the parameter coefficient are constrained by the tuning parameter t —known as the penalty. When $t = 0$, parameter coefficient β is the linear regression estimate, and as t increases, it shrinks the parameter coefficient. Lasso regularization sets to zero the parameters that do not impact the dependent variable. Thus, lasso regularization performs variable selection as it “throws out” irrelevant variables.

Wheeler (2009) tested the effectiveness of the Least Absolute Shrinkage and Selection Operator (LASSO) applied to geographically weighted regression to manage multicollinearity in coefficients against non-regularized GWR and ridge-regularized GWR. He detected multicollinearity in median income and housing rates when applying non-regularized GWR to a Columbus, Ohio crime dataset. Wheeler found that the GWR model with localized LASSO outperformed GWR when applied to the crime dataset. However, he also found that the localized LASSO overshrinks regression coefficients

²⁹ (Tibshirani, 1996)

and therefore, recommends it only be used to predict the response variable. In this research context, LASSO would be applied to the regression coefficients, necessitating the application of global, rather than local LASSO. However, while globalized LASSO was shown to be more accurate, it is not a viable method for large datasets, as computational demands could exceed computer memory capacity. This dissertation research involved a library of large datasets, and thus would not have been viable had global LASSO been utilized.

Elastic net regularization was developed by Zou and Hastie (2005) to address limitations of the Lasso regularization technique. Lasso saturates when accommodating a high number of predictor variables. It also eliminates multicollinearity by eliminating all but one of the collinear variables. Zou and Hastie demonstrate the effectiveness of their method using a genetic research case study, where the sample size is small (less than 100) with respect to the variable array (thousands of genes). Gene correlations are often present, as many share the same biological pathway. Elastic net will select variable clusters if one variable in that cluster is deemed relevant. The research data in this dissertation included many variables that would behave as clusters. For instance, crime statistics were available for burglaries, car thefts, and robberies. These crime statistics increased and decreased together over time and space. Lasso would eliminate all but one of the variables, and OLS would require the researcher to omit two of the three variables to eliminate the multicollinearity. This requirement implies that the inclusion of all three variables would contribute nothing to the model, but this might not be the case. It's possible that nuances in behavior could be lost. Elastic net regularization allows the

incorporation of collinear variables, grouping the coefficients into an elastic net framework.

Elastic net is a hybridization of L1 and L2 norms. Whereas for ordinary least squares, the regularization parameter α is zero and for Lasso α is equal to one, α values for elastic net lie between zero and one. This regularization technique could be the preferred alternative if ordinary least squares is too rigid and lasso is judged as too relaxed for the problem set.

Related research has generally been focused on identifying risks so that potential targets can be protected. It is hoped that this model will improve security through prevention, not only by providing insights into drivers and mitigating factors to better identify solutions, but also by helping to identify which questions should be asked to implement a solution more effectively.

For instance, the model developed in this research can direct stakeholders' attention to areas that might require changes in policy implementation, resource allocation—and depending on the behavior and level of coefficient impact—provide insight into what changes might be most effective to minimize an area's vulnerability to security threats associated with radicalization. For example, if municipal representation manifests as a negatively correlated coefficient in some regions but shows little correlation in others, resources aimed at encouraging group members to enter politics could be concentrated

in those regions most affected by that coefficient. Moreover, this targeted approach can be fiscally prudent because resources would not be expended where they are unlikely to yield significant dividends.

Furthermore, this model can reveal key questions that might otherwise not be asked. For example, starkly contrasting coefficient behavior in different regions might reveal an unusual dynamic, and variables can have mutual effects that vary as conditions change. These contrasting behaviors are evident in the model results, but often this type of dynamic is active for years and goes unnoticed or is only vaguely apparent outside the context of the model. There are many possible explanations for contrasting coefficient behavior. One possibility could be the impact of additional variables that are present in some of the regions, yet are absent or play a muted role in other areas.

Another possibility might rest with the driver sustaining the variable. For instance, consider the hypothetical case of country whose high-income coefficients exact contrasting effects in different districts. One of the high-income districts might owe its financial health to a resident population that resides outside the country, but these seasonal residents purchase property because they spend much of their time in the district and have the wealth to do so. These part-time residents affect the demographic, but they might have little effect on local industries, commerce, and have limited interaction with natives. The impact of these high-income residents could be quite different from that of residents of another district who gained their wealth through local entrepreneurial endeavors or through professional training. This group might raise their children in the area, and might perhaps identify more closely with and have a deeper commitment to the area in which they reside. The model output of this particular case

might reveal that an emotional stake in the area, or the attitude of permanency that often accompanies those who “lay down stakes” to raise families, reveals itself as an unusually strong variable that is the true coefficient—not high-income. Thus, in this example, the model enabled analysis that cast aside a less relevant question, “what is the effect of high-income on a district’s vulnerability to harboring a safe-house in country X?”, and shed light on the question that should actually be asked, which is “how can we deepen commitment to a space and associate that commitment with identity?” I offer this exposition as a hypothetical example to illustrate how the model can facilitate the development of questions that might be more effective catalysts for identifying solutions for at-risk areas. Finally, this research has laid the foundation for future refinements and applications of this approach. Thus, it is hoped that going forward this work will deepen our understanding of multivariate interdependencies and how these differ among countries in the course of further refinements and applications of this model. This research should, on a practical level, focus attention on vulnerable areas before that area becomes destabilized, and serve as a guide to facilitate resource allocation that enhances quality-of-life for communities.

Note on Map classification of variables and GWR coefficients.

With the exception of maps displaying Local R² values, classes are divided using the Natural Jenks Optimization Method. This classification technique was developed by George Frederick Jenks and is often used to display data on maps. This method finds natural breaks in the data by comparing the sum of the squared means deviations to define the classes such that the average deviation from the class mean is minimized, but the deviation between the classes is maximized. (Brunsdon, Fotheringham, & Charlton, 1998)

This does not constitute an official release of United States Government (USG) information. All statements of fact, opinion, or analysis expressed are those of the author and do not reflect the official positions or views of the USG. Nothing in the contents should be construed as asserting or implying U.S. Government authentication of information or endorsement of the author's views. This material has been reviewed for classification and compliance with legal obligations.

Bibliography

(2008). Retrieved from State Department Bureau of European and Eurasian Affairs: <http://www.state.gov/r/pa/ei/bgn/3842.htm>

Adida, C., Laitin, D., & Valfort, M.-A. (2009). Integration into Europe: Identifying a Muslim Effect. For presentation at the December 11-12, 2009 meeting of the Working Group on African Political Economy.

Adida, C., Laitin, D., & Valfort, M.-A. (2012). Muslims in France: Identifying a Discriminatory Equilibrium. Institute for the Study of Labor.

Al Arabiya. (2012, November 21). Against all odds, 'Mosque for gays' to open in France. al Arabiya.

Associated Press. (2004, January 19). French Town Notorious for Islamic Radicalism. Associated Press.

Azam, J.-P. (2010). Why Suicide-Terrorists Get Educated and What to Do About It. Toulouse University, Toulouse School of Economics.

BBC. (2010, October 19). France Roma Expulsions. BBC.

Belmssous, H. (2002, November 14). Price to pay for 'island of privilege': fenced off in France. *Le Monde diplomatique*.

Berrebi, C. (2007, December). Evidence About the Link Between Education Poverty and Terrorism Among Palestinians. *Peace Economics, Peace Science and Public Policy*, 13(1), 1554-8597.

Bhui, K. (2014). Is Violent Radicalisation Associated with Poverty, Migration, Poor Self-Reported Health and Common Mental Disorders? (X. J. Yong-hui Dang, Ed.) 9(3).

Boekhout van Solinge, T. (1995). Cannabis in France. Cannabis Policy, Criminal Law and Human Rights Conference. Germany: Bremen University.

Brown, D., Dalton, J., & Hoyle, H. (2004). Spatial Forecast Methods for Terrorist Events in Urban Environments. Presentation at the Second Symposium on Intelligence and Security Informatics.

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression--modelling spatial non-stationarity. *The Statistician*, 47(3), 431-443.

Camilleri, R. (2012). Impact of Counter-Terrorism on Communities: France. Background Report, Institute for Strategic Dialogue.

Caron, P.-L. (2015, July 27). Police Seize Nearly Six Tons of Hashish in Major Drug Bust in France. *Vice News*. Retrieved from <https://news.vice.com/article/police-seize-nearly-six-tons-of-hashish-in-major-drug-bust-in-france>

Chang, A. (n.d.). Table of Sample Size Requirement for Multiple Regression. Retrieved from StatsToDo: http://www.statstodo.com/SSizMReg_Tab.php

Chiriac, M. (2012, September 11). France, EU, Seek Action on Roma from Romania. Balkan Insight.

Dutton, P. (2007, August 11). France's Model Healthcare System. The Boston Globe.

Evry. (2004, February 4). The war of the headscarves. The Economist.

Fields, R. M., Elbedour, S., & Hein, A. (2002). The Palestinian suicide bomber. In C. ed. Stout, The Psychology of Terrorism: Clinical Aspects and Responses (pp. 193-223). Westport: Praeger Publishers.

Finney, N., & Simpson, L. (2006). Place of residence of Muslims charged with terrorism under UK legislation. NonTechnical Summary Report, The University of Manchester, Cathie Marsh Centre for Census and Survey Research.

Francica, J. (2007, September 5). Assigning Risk to Targets of Terrorism - A Location Intelligent Approach. Directions Magazine.

Gennip, Y. v., Hunter, B., Ahn, R., Elliott, P., Luh, K., Halvorson, M., et al. (2012, November). Community detection using spectral clustering on sparse geosocial data. SIAM Journal on Applied Mathematics, 73(1), 67-83.

Government of France. (2007). Implementation of the International Covenant on Economic, Social, and Cultural Rights. United Nations Economic and Social Council. New York: Government of France.

Groff, E. R. (2008). Adding the Temporal and Spatial Aspects of Routine Activities: A Further Test of Routine Activity Theory. Security Journal(21), 95-116.

Gurr, T. R. (1970). Why Men Rebel. Princeton: Princeton University Press.

Hassan, N. (2001, November 19). An Arsenal of Believers. The New Yorker.

Henry, M., & Dieusaert, P. (2014). Onzus Infos--L'emploi des femmes dans les zones urbaines sensibles. Observatoire National Des Zones Urbaines Sensibles. Département études statistiques et systèmes d'information du SG.

Hertog, S. (2008, September 1). Extremist Engineers. (S. Karlin, Interviewer) IEEE Spectrum.

Hinnant, L. (2012, October 12). 1 dead, 10 arrested in anti-terror sweep in France. Associated Press.

Holman, T. (2015, January 15). The Clear Banner: French Foreign Fighters in Iraq 2003-2008. Jihadology.net.

Integrated Regional Information Networks. (2013, May 9). Countering the radicalization of Kenya's youth. IRIN Humanitarian News and Analysis.

Ireland, D. (2005, November 06). WHY IS FRANCE BURNING? The rebellion of a lost generation.

Jones, J. D., & Wiseman, R. J. (2006, September 8). Community Policing in Europe: Structure and Best Practices-Sweden, France, Germany. Retrieved 2015, from Los Angeles Community Policing.

Knofczynski, G., & Mundfrom, D. (2008, June). Sample Sizes When Using Multiple Linear Regression for Prediction. *Educational and Psychological Measurement*, 68(3), 431-442.

Leonard, T. (2015, March). France Changing--What this Weeks Elections Mean for France and the World. Retrieved September 7, 2015, from Values & Capitalism: <http://www.valuesandcapitalism.com/france-changing-what-this-weeks-elections-mean-for-france-and-the-world/>

Macy, M., & Miller, R. (2001). From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, 28, 143-166.

Marcel, M. (2009). Constructing Mosques: The Governance of Islam in France and the Netherlands. Amsterdam Institute for Social Science Research (AISSR).

Marshall, J. (2008, July 20). France: More Super-Campuses Announced. *University World News* (37).

McCauley, C., & Moskalenko, S. (2008). Mechanisms of Political Radicalization: Pathways Toward Terrorism. *Terrorism and Political Violence*, 20, 415-433.

Mili, H. (2006, January 12). Tangled Webs: Terrorist and Organized Crime Groups. *Terrorism Monitor*, 4(1).

Niebuhr, R. (1960). *Moral Man and Immoral Society*. New York: Scribner.

Open Source Center. (2009). *Afghanistan-Geospatial Analysis Reveals Patterns in Terrorist Incidents 2004-2008*. Washington, DC: Open Source Center.

Parrish, K. (2012, March 28). Link Grows Between Terrorism, Organized Crime, Officials Say. Retrieved 2015, from Defense.gov: <http://www.defense.gov/news/newsarticle.aspx?id=67721>

Peterson, F., & Fray, P. (2004, January 16). Brigitte Denies Terrorist Links in Letter. *Sydney Morning Herald*.

Phelan, J. (2012, December 4). Police Shooting: Police Arrest 2 Suspected Accomplices. *Global Post*.

Pizzi, M. (2014, March 19). Study finds unexpected risk factors for violent radicalization. al-Jazeera America.

Richardson, C. (2011). Relative Deprivation Theory in Terrorism: A Study of Higher Education and Unemployment as Predictors of Terrorism. New York: New York University Politics Department.

Russell, C. A., & Miller, B. H. (1977). Profile of a Terrorist. *Terrorism*, 1(1).

Sageman, M. (2004). *Understanding Terror Networks*. Philadelphia: University of Pennsylvania Press.

Shafiq, M. N., & Sinno, A. H. (2009, August). Education Income and Support for Suicide Bombings-Evidence from Six Muslim Countries.

Shelley, L., Picarelli, J. T., Irby, A., Hart, D. M., Craig-Hart, P. A., Williams, P., et al. (2005, June 23). Methods and motives: exploring links between transnational organized crime & international terrorism. National Institute of Justice, Office of Justice Programs, U.S. Department of Justice.

Shi, H., Zhang, L., & Liu, J. (2005). Incorporating Geographic Space and Attribute Space into the Kernel Weighting Function of Geographically Weighted Regression. 8th International Conference on GeoComputation. Ann Arbor: Geocomputation Conference.

Shi, H., Zhang, L., & Liu, J. (2006, April). A new spatial-attribute weighting function for geographically weighted regression. *Canadian Journal of Forest Research*.

Simon, P. (2012). French national identity and integration: who belongs to the national community? Transatlantic Council on Migration. Washington, DC: Migration Policy Institute.

Simon, P., Beauchemin, C., & Hamelle, C. (2010). Trajectories and Origins: Survey on Population Diversity in France. Retrieved from Institut National d'Études Démographiques (INED): <http://EconPapers.repec.org/RePEc:ide:wpaper:168.1>

Smith, B. L., Damphousse, K. R., & Roberts, P. (2006). Pre-Incident Indicators of Terrorist Incidents: The Identification of Behavioral, Geographic, and Temporal Patterns of Preparatory Conduct. Washington, DC: National Criminal Justice Reference Service.

Snijders, T. A., Steglich, C. E., & Schweinberger, M. (2007). Modeling the Co-evolution of Networks and Behavior. In K. Eds. van Montfort, H. Oud, & A. Satorra, *Longitudinal models in the behavioral and related sciences* (pp. 41-71). Mahwah, New Jersey: Lawrence Erlbaum.

State Department. (2015). U.S. Relations with France. State Department Bureau of European and Eurasian Affairs.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267-288.

Tlemçani, R. (1997, March). Islam in France: The French Have Themselves to Blame. *Middle East Quarterly* , 4(1), 31-38.

Townsend, E. (2007). Suicide Terrorists: Are they suicidal? *Suicide and Life Threatening Behavior*, 37(1), 35-49.

Victor, B. (2003, October 17). Raising a Generation of Martyrs. *Chicago Tribune*.

Victor, B. (2004). *Army of Roses: Inside the World of Palestinian Women Suicide Bombers*. Robinson Publishing.

Viorst, M. (1996, September/October). The Muslims of France: Islam Abroad. *Foreign Affairs Journal*.

Way, D. (2003). Targeting Terrorists: Can GIS Lead the Attack? . *GEO World*.

Wheeler, D. C. (2009). Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. *Environment and Planning A*, 41, 722-742.

Willis, H. H., LaTourrette, T., Kelly, T. K., Hickey, S., & Neill, S. (2007). *Risk Modeling for Intelligence Analysis and Infrastructure Protection*. RAND Corporation. Center for Terrorism Risk Management Policy.

Woo, G. (2003, April 22). Evolution of Terrorist Risk Modeling. *Journal of Reinsurance*,7.

Würger, T. (2014, August 13). Suras and Tolerance: Meet Europe's Gay Imam. *Spiegel Online*.

Yildirim, J., Ocal, N., & Korucu, N. (2013). Analysing the Determinants of Terrorism in Turkey using Geographically Weighted Regression. *Defence and Peace Economics*, 24(3), 195-209.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society B*, 67(2), 302-320.

VITA

ELYKTRA CLAIRE EISMAN

Born, Miami Beach, Florida

B.S., Electrical Engineering
Florida International University
Miami, Florida

2009 Research Assistant in GIS/Remote Sensing
Disaster Risk Reduction Center
Florida International University
Miami, Florida

2011 M.A., International Studies
Florida International University
Miami, Florida

Research Assistant for GIS/Remote Sensing
Disaster Risk Reduction Center
Florida International University
Miami, Florida

Graduate Certificate in National Security Studies
Florida International University
Miami, Florida

2011 - 2015 Doctoral Candidate
Florida International University
Miami, Florida

PUBLICATIONS AND PRESENTATIONS

Eisman, Elyktra. *The Role of GIS in DRR: Classifying and improving GIS capabilities in small towns and rural areas*. working paper. Florida International University. Miami: Disaster Risk Reduction in the Americas Latin American and Caribbean Center, 2009.

--. "Online News Media Impact on Disaster Response: October 26, 2010 Indonesian Disaster Case Study." *The International Journal of the Constructed Environment* 2.2 (2012): 143-178.