Social Networks 36 (2014) 23–39

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Comparative configurational analysis as a two-mode network problem: A study of terrorist group engagement in the drug trade

Ronald L. Breiger[a,*], Eric Schoon[a], David Melamed[b], Victor Asal[c], R. Karl Rethemeyer[d]

[a] School of Sociology, University of Arizona, Tucson, AZ 85721, United States
[b] Department of Sociology, University of South Carolina, Columbia, SC 29208, United States
[c] Department of Political Science, Rockefeller College of Public Affairs and Policy, University at Albany, Albany, NY 12222, United States
[d] Department of Public Administration and Policy, Rockefeller College of Public Affairs and Policy, University at Albany, Albany, NY 12222, United States

ARTICLE INFO

Keywords:
Duality
Two-mode network analysis
Terrorism
Illicit economies

ABSTRACT

We generalize a form of two-mode network analysis to make it applicable to a cases-by-variables data format, and apply our approach for the study of terrorist group engagement in the drug trade, emphasizing the implications of our approach for policy in a study of 395 terrorist organizations. Based on the organizations' levels of resources, network connectivity to other groups, ideological emphasis, and participation in multiple illicit economies, we identify several distinctive configurations of factors that lead to multiple types of drug activity. We also demonstrate a technique for assessing sampling variability in configurational models.

1. Introduction

The nexus of drugs and terrorism as a problem in political networks has been addressed from several different angles of research. First, because terrorist groups in general often use violence and intimidation to challenge, compete with, attack, or provide alternatives to the political authority of states, analysts view terrorism as a distinctive and consequential form of political violence, and are continuing to shape network analysis methods to study it (Perliger and Pedahzur, 2011). Second, different kinds of network structure and context carry differing implications for how terrorist networks strive for resilience, including when and how such groups may combine drug-related and other forms of illicit activity with nationalist or other political motives (Bakker et al., 2012). These varying approaches to resilience in turn have implications for counterterrorism policies (Milward and Raab, 2006, 2009). Roberts and Everton (2011; Everton, 2012) set forth an agenda along with much specific guidance and examples for embedding social network analysis techniques and interpretation of their output within larger strategic and theoretical frameworks for countering illegal and covert networks. Third, as emphasized in a hotly-debated recent Brookings Institution study, illicit economies (with drugs as the paradigmatic case, but also including trade in chemical and biological weapons components, human beings, conflict diamonds, and other illicit goods and services) provide a specific form of political capital to terrorist organizations. By "protecting the illicit economy," terrorists "protect the local population's livelihood from government efforts to suppress it" (Felbab-Brown, 2010: 17). Implications for public policy deriving from Felbab-Brown's political capital model of illicit economies include the controversial recommendation of less emphasis on eradicating drug crop cultivation per se and more on local government efforts to gain legitimacy with the local population (Felbab-Brown, 2010: 156–184).

What factors are associated with the participation of a relatively small number of terrorist groups in the drug trade, while insulating most others from this activity? How does drug activity relate to the wresting of territory from state control, ethnic grievances, and the pursuit of unconventional weapons (chemical, biological, radiological, or nuclear), among other activities and organizational attributes? We generalize a form of two-mode network analysis to address these and related questions in a study making use of open-source data on 395 terrorist groups in the period 1998–2005. In the process, we engage with the first two perspectives listed above on terrorism as a problem in political networks, and we also put forward a distinction relevant to Felbab-Brown's political capital model according to which, in our analysis, there are multiple logics moving terrorist groups toward drug activities, suggesting (again in our view) the possibility of different policy implications in different contexts.

Important studies of terrorist connections have been conducted on full network data ("who-to-whom" and "who-to-what") derived from open sources (e.g., Everton, 2012; Krebs, 2001; Pedahzur and Perliger, 2009; Roberts and Everton, 2011; Rodriguez, 2005). Nonetheless, in many situations information on the ties among terrorists is notoriously "incomplete, inaccurate or simply not available" (Tsvetovat and Carley, 2005; see also Hayden, 2009;

* Corresponding author. Tel.: +1 520 621 3524; fax: +1 520 621 9875.
E-mail address: breiger@arizona.edu (R.L. Breiger).

0378-8733/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.socnet.2013.04.002

Sparrow, 1991). One highly productive reaction has been to focus on computational modeling in order to understand behavior on the basis of simulated terrorist networks (e.g., Tsvetovat and Carley, 2005). In this paper we pursue a different strategy, making use of database information on actual groups and some of their known behaviors and attributes. As Perliger and Pedahzur (2011) point out, there has been "a striking increase in efforts and resources invested in data collection" on terrorist groups in recent years by academic and governmental agencies, such as the open-source, publicly available databases maintained at the START Center at the University of Maryland, resulting in the present availability of "high-resolution" information (see also Hayden, 2009). We believe that the community of social network analysts has important methods to offer that can be further developed to apply to analyses of data on terrorism of the sort that is contained in existing, public databases derived from open-source data.

1.1. Statement of objectives

This paper aims to make contributions that are both methodological and substantive. As to method, we begin by noting the importance to network analysis of two-mode (and, more generally, multi-mode) analysis of ties of affiliation connecting actors at different levels of structure (such as persons and groups). We move on to propose a substantial generalization, to make two-mode analysis relevant to the kinds of data table that typically underlie regression analyses and other methodologies that operate on cases-by-variables formats, such as Qualitative Comparative Analysis (e.g., Ragin, 2008), on which we focus here. Thus, techniques for the study of affiliations in two-mode networks can be generalized so as to be relevant to research problems that are often approached currently via configurational or regression modeling.

We demonstrate our methodological approach by showing that it offers new substantive insight. We explore correlates of participation in the drug trade on the part of terrorist groups. Our network-inspired two-mode thinking pushes us to look for multiple typologies (just as analysts have discovered multiple nexuses of persons and events in the "Southern Women" dataset of Davis, Gardner, and Gardner, for example [see Freeman, 2003, for extensive review and analysis]). Felbab-Brown (2010) puts forward in quite general terms her influential critique of eradication as a primary counter-narcotics policy. However, we emphasize that different cases manifest different sets of contingencies (see also DuPée, 2010 on multiple roles with regard to illicit drug activities, and George and Bennett, 2005, on case studies and theory development in the social sciences). Specifically, we find variety in types of involvement by groups engaged in drug activities, some groups exhibiting strong control of territory and tending to exploit directly the process of cultivation, while other groups do not control territory and are in many cases far removed from the drug growing fields. We consider the possibility that the eradication-centered policies that Felbab-Brown criticizes may in fact be effective when applied to the latter set of terrorist groups and their drug involvements. Thus, we provide a new methodological opening to discovering multiple types of admixtures of cases and variables that build on two-mode reasoning and that can aid in the study of policy alternatives. Our analysis does not include all possible variables required for a comprehensive study of drugs and terrorism, nor do we present a substantial theory of that subject. Instead, we show the potential of our approach to help researchers improve their ability to assess terrorist groups' degree of existing involvement in the drug trade.

2. Background

The two subsections focus in turn on introducing the generalization and expansion we propose of two-mode network analysis, and on providing brief and focused reference to the research literature that features case studies and substantive analyses at the intersection of illicit economies and violent non-state actors (VNSAs). From the existing research literature we identify behaviors and attributes that are of particular interest in linking terrorist organizations with participation in the drug trade.

2.1. Two-mode analysis of actors and attributes

The key concept of social network analysis that we further develop in this paper is duality, which has long been defined as the "turning inside-out" of a network at one level (say, pairs of groups as actors, connected by the people who belong to both groups in the pair) such that the vertices at one level become the edges at another (for example, pairs of persons as actors who are connected by the groups with which they jointly affiliate; Breiger, 1974). The concept of duality has been extended to multiple levels of affiliation (multimode analysis, with Fararo and Doreian, 1984, pioneering tripartite analysis; see also Carley, 2003, Cornwell et al., 2003, for major developments) and to sets of social networks at different levels (such as citations among French cancer researchers' publications, and mobility of researchers from one laboratory to another) connected by membership ties (affiliations of researchers with laboratories; Lazega et al., 2008), and the basic idea has been incorporated within the powerful ERGM family of statistical models for networks (Wang et al., 2009).

In this paper we develop an application of the duality concept that is relevant to database information on terrorist groups and their attributes and behaviors. Our approach to information in databases builds on Burt's (1983) fundamental insight that,

> "A connection exists between the usual concept of an actor's network position in social structure and combinations of attributes defining statuses in that social structure. Survey data on a randomly sampled respondent can be used to describe the relation pattern defining his 'ersatz' network position in the social structure from which he has been drawn."

By the term "ersatz" Burt means to imply that an actor's position (whether that actor is a man or woman or an organization) based on "combinations of attributes" can serve as a potentially useful artificial substitute for the actor's position in social structure. In this paper we have no need to think of an actor's location in an intersecting lattice of variables as a substitute for a social network among actors, because we consider the actor's position in a space of configurations of variables to be of interest in its own right. It may be difficult to obtain data on network ties among terrorist groups, but the relative similarities of terrorist groups to one another within a space of attributes, behaviors, and proclivities is an interesting generalized network to study in its own right, even though it exists on a different plane from direct connections of communication or resource flow among the same organizations. Similar insight has motivated network analysts to develop algebraic representations for beliefs and attitudes on the basis of survey data (Martin and Wiley, 2000), to study "symbolic networks" linking actors to political symbols (Ansell, 1997; Cunningham et al., 2010), to develop Bourdieu's concept of field theory with reference to correspondence analysis of data on group attributes and behaviors (de Nooy, 2003), to recast regression modeling as a problem in network analysis (Breiger et al., 2011; Melamed et al., 2013), and to analyze actor-by-event data matrices by using a form of Boolean analysis that is very similar to the QCA approach of Charles Ragin (Schweizer, 1996).

Ragin's set-theoretic methods of Qualitative Comparative Analysis (QCA; Ragin, 2000, 2008) are widely known in sociology, political science, and related disciplines as providing means for analysis of cases-by-variables datasets that are "case-oriented" and that speak to a different set of questions than those typically pursued via multiple regression methods and their many generalizations (see also Mahoney and Goertz, 2006). As Rubinson and Ragin (2007) describe the difference,

> "It is not that case-oriented researchers dismiss variables, but rather that they perceive that it is not variables but cases that have relationships with one another. The variable-oriented researcher shows, for example, that poverty is correlated with democracy. The case-oriented researcher observed that criminals tend to be poor, especially the ones that get caught, and that economically developed countries tend to be democratic. Although subtle, this distinction entails fundamentally different views of social phenomena. Where variable-oriented researchers view the social world as a manifestation of the myriad relationships among variables, case-oriented researchers see many different kinds or sets of cases."

As with the study of comparative politics more generally (Siegel, 2011), the comparative study of the behavior of terrorist groups is in large part the study of the role that context plays in structuring behavior. Ragin's QCA approach provides a well-developed and widely-used means to map that context.

In this paper we show that it is productive for some purposes to think of configurational methods, including Ragin's QCA as well as the complementary techniques that we will apply, as a form of two-mode network analysis, one that emphasizes what Breiger (2009) has termed the duality of cases and variables. Ragin has insightfully insisted that cases are configurations of variables, and we argue that it is just as important to see variables as configurations of the cases that comprise them (which we believe typically remains an implicit assumption in QCA). While this assertion may, on first glance, be unintuitive, consider the following: when a researcher selects a variable for use in analysis, there is typically an embedded assumption that the variable represents, or is indicative of, a social phenomenon. Our assertion that variables should be viewed as configurations of cases is rooted in the idea that the way a given variable represents a phenomenon is not intrinsic to the variable itself, but rather is rooted in the cases that the variable is meant to represent. For example, while level of income is a standard indicator, what constitutes "high" versus "low" income will numerically be very different if you are studying graduate students as opposed to hedge fund managers. As such, while the way a variable is measured may be determined entirely independent of a given case, the empirical meaning of the variable is constituted by the cases it is used to represent.

With reference to the analysis of database information on terrorist groups, we illustrate the application of a form of two-mode analysis that captures several of the key features of QCA while emphasizing the duality of cases and variables. In addition, we provide measures of the effect of sampling variability on estimates of key quantities associated with the configurational models that we discover in our dataset. The main analytic machinery that we use is a form of dimensional analysis ("barycentric correspondence analysis") that is particularly well-suited for the set-theoretic work at the core of QCA, and that leads to a form of visualization that productively complements QCA analyses.

2.2. *Terrorist involvement in the drug trade*

A decade ago influential analysts described considerable differences between drug and terror groups regarding motivations and operations (Kenney, 2003), but that boundary appears to be increasingly blurring, with major implications for terrorist capabilities (Murch and Tamsett, 2009: 250). Research focused on the intersection of illicit economies and terrorist actors has been evolving at least since the late 1970s, when violent non-state groups in South America—most notably the Revolutionary Armed Forces of Colombia (FARC) and Sendero Luminoso (Shining Path) of Peru—began to participate in the worldwide the drug economy to finance their military operations (Makarenko, 2004; Felbab-Brown, 2010). With the end of the Cold War, and subsequently the post-9/11 success of governments in limiting existing sources of funding, the relatively small proportion of violent non-state actors engaged in the drug trade (Steinitz, 2002) began to increase. As a result, the collaboration and mutual integration of organized crime and terrorism has become a central concern for both scholars and policy makers (United Nations General Assembly, 1999; Cilluffo, 2000; Makarenko, 2004). In the past, drug traffickers were a secondary concern for intelligence on terrorism; now, this subject matter is often of primary interest (Treverton, 2009).

While the risks associated with VNSAs' involvement in illicit economies are significant, the proportion of known occurrences remains comparatively small. As a result, the common challenges associated with drawing general conclusions from case studies of infrequent social phenomena become quite real in this field of research (see Harding et al., 2002). Case studies of violent groups engaged in illicit activities have developed invaluable insights for understanding this social problem (e.g., Bakker et al., 2012; DuPée, 2010; Felbab-Brown, 2010; Milward and Raab, 2006; Morselli et al., 2007; Raab and Milward, 2003; Roth and Sever, 2007). We seek to benefit from these insights and retain some of the interpretative features of qualitative analysis while also providing for a wider panorama of comparative possibilities (hundreds of cases rather than dozens or fewer) that are made possible by open-source databases.

The literature referenced above emphasizes a wide variety of attributes and behaviors of organizations engaged in violent activities that are associated with participation in the drug trade. This research provided the basis for the measures included in our analysis. While theory was the primary guide of our variable selection, we also had to address the critical logistical constraint of finding reliable data that was comparable across all cases. As should be expected of any study that uses multi-national quantitative data, this constraint prevented us from accounting for the full spectrum of predictive factors that have been identified through in-depth case research of drug smuggling organizations. However, the factors selected meet the necessary standards of reliability and comparability, and are representative of key points of consensus across this literature on conditions that predict a terrorist group's engagement in the illicit drug economy. The factors are these.

2.2.1. *Resources*

An important perspective that Milward and Raab (2006) derive from their case studies is that it is not variables, but trade-offs, that are most relevant to the analysis of organizations that are covert and illegal. These authors discuss in particular the trade-off between integration and differentiation. Integration is the ability of a "dark network" to grow in size through recruitment, for example, and the ability to hold territory—features that are paradigmatically relevant to the organizations illicitly involved in the Colombian drug trade, which is one of the principal case studies in their 2006 paper (see also the expanded discussion of the role of terrorist groups in the Colombian drug trade, in Bakker et al., 2012). Differentiation on the other hand involves a group becoming more horizontal and segmented (less centralized) and less visible in response to "shocks" from government authorities. Capacity for action (related to integration) might be traded off to increase persistence (as a dispersed form may at times be necessary for survival), as reflected in a major proposition defended by Milward and Raab. A related general proposition is that "a variety of resources—territory, technology, finances, weapons, and law—serve as necessary but not sufficient conditions" for the development of dark networks (Milward and Raab, 2006).

2.2.2. Connectivity

An extensive literature demonstrates the value of network connections for social movements (Diani and McAdam, 2003) and for covert networks (Arquilla and Ronfeldt, 2001). Both through cooperative and competitive alliances, terrorist groups benefit from ties to other groups (Asal and Rethemeyer, 2009; Raab and Milward, 2003). Being involved in lucrative drug distribution networks often implies having far-flung network ties, even stipulating the need for secrecy. Makarenko (2004) argues that interaction and cooperation among drug smugglers and terrorist groups results in these organizations "learning from one another, and adapting to each other's successes and failures."

2.2.3. Knowledge and skills

Felbab-Brown (2010: 190) notes that, "in the initial stages, belligerent groups rarely have the resources and know-how to set up a large illicit economy." Groups with the knowledge, skills, materials, and abilities to engage in one type of illicit trade are likely to have improved capabilities for engaging in another type of illicit trade. The same organizations that smuggle drugs might therefore engage in trafficking diamonds, humans, or substances or parts that can be used in chemical or radiological weapons (Asal and Rethemeyer, 2009; Milward and Raab, 2006; Raab and Milward, 2003; for a popularized account, Naim, 2005). A staff member of the National Academies (testifying in his own behalf) told a subcommittee of the U.S. Congress that "drug networks should probably be of greatest concern in considering the future of international smuggling of material for dirty bombs" (Schweitzer, 2005).

2.2.4. Ethnic ideology

In a review of political science literature on Third World security, Thomas (2003: 217) argued that

> "The illicit drug trade now extends beyond drug cartels seeking commercial profit. Illegal drugs have become a lucrative source of income for dissident ethnic groups who use the proceeds to purchase arms and equipment for conducting violent campaigns for secession, or for other domestic or international political causes."

The author was referring to groups such as the Kosovo Liberation Army in the former Yugoslavia and the Liberation Tigers of Tamil Ealam (Tamil Tigers) in Sri Lanka as of the time of writing. Presumably the bonds among members of the same ethnic group, especially when they confront what they perceive to be threats to their very existence, foster the kind of solidarity that warrants obtaining resources, so to speak, by any means necessary (Paoli and Reuter, 2008).

In addition to solidarity, a shared ideological component that features ethnic bonds may also be expected to lead to increased trust, an idea that features strongly in social capital theory (Coleman, 1988) and in identity theory (Tajfel, 1981). The interplay of ethnicity and drug supply networks in the South Asian communities of Britain is studied in Ruggiero and Khan (2006).

2.2.5. Other factors; other questions

We emphasize that this review, and our research project more generally, has not captured all the variables needed to fully understand terrorist involvement in the drug trade. However, based on a review of the extent literature, we believe that the factors on which we have data and on which we focus are sufficient to enable us to demonstrate the potential of our approach as one means for enhancing rigorous analysis of an important policy topic. The four factors reviewed above focus more on capabilities than on intent, but we do not want to give the impression that terror groups become involved with drugs simply because they can. Other key factors potentially at work might over time be included in a more robust version of the approach we are about to present. For example, might governments' success in reducing terrorist funding prompt some groups, partly out of desperation, to turn to illicit economic activity? If in consequence some terror organizations increasingly involve themselves in the drug trade, will this raise their operational profiles and make them more vulnerable to counterterror efforts? What is the role of geography (including locales where drugs can and cannot be grown) and how does that interplay with control over territory (the latter being a variable we include in our study)? What are the unintended consequences of states' counter-narcotics and counterterrorism policies (Felbab-Brown, 2010)? What is the social ecology within which various kinds of groups (cartels, terrorists, insurgents, traffickers, warlords, local government officials) adopt as well as adapt or morph identities across a range of domains? These questions and concerns help us to emphasize that the interrelations of the variables that we portray in this study are one part of a much larger picture.

3. Data and methods

In spite of the valuable contributions of the research we reviewed in the previous section, there is a relative dearth of formal models that allow analysts systematically to locate a panorama of violent non-state organizations within the frameworks presented. We now proceed to demonstrate a modeling framework to leverage analysis on the basis of a two-mode network analysis. This work is inspired by the highly innovative comparative configurational modeling of Ragin's QCA, and in particular by our interest in bringing to the forefront the duality of cases and variables that we believe to be a key, if often implicit, assumption of Ragin's approach.

3.1. Data

To demonstrate the applicability of this methodological approach to the substantive questions outlined above, we use data from version 1 of the Big Allied and Dangerous (BAAD-1) database. This data is housed and maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START)—a Center of Excellence of the U.S. Department of Homeland Security—and is currently the most extensive database publicly available on terrorist activities and attributes that takes organizations as the units of analysis. Compiled from open sources, these data have been widely used in research ranging from the study of the lethality of political violence (Asal and Rethemeyer, 2008) to terrorist groups' plots to use, as well as uses of, chemical, biological, and radiological weapons (Asal et al., 2012). Data from the lethality study are publicly available (Asal et al., 2009a). In this study we use several additional variables from the same database. This cross-sectional dataset contains information on 395 terrorist organizations that were known to be active between the years of 1998 and 2005. The data matrix is in a cases (organizations) by variables (properties) format well suited for the application of two-mode network analysis. Information on the genesis of the database is provided in Asal and Rethemeyer (2008, 2009; Asal et al., 2009a), and more detailed information on the coding and data collection procedures can be found in Appendix A.

The outcome variable, an organization's participation in the drug trade, was coded zero if we found no evidence that an organization trafficked or produced drugs in the period from 1998 to 2005 and a one otherwise. This coding was done under the supervision of two of us using newspapers, online resources and available books that detailed the activity of the 395 terrorist groups in the sample (for more information, please see Appendix A). Thirty-five of the 395 organizations (8.9%) were coded as participating in drug trafficking or production.[1] We will often use the letter **Y** to denote this binary outcome variable.

We selected five properties (correlates potentially useful for locating cases within a space of properties) that reflect the attributes and behaviors emphasized above in our review of the case-study and substantive literature on terrorism and participation in the drug trade (Section 2.2.). To represent resources, we use the variables of organizational size (**S**) and strong control of territory (**T**). Data on size was collected initially from the Terrorism Knowledge Base (TKB). To validate this data, experts at the Monterey Terrorism Research and Education Program of the Monterey Institute for International Studies (MIIS) and members of the U.S. intelligence community were provided with a complete list of organizations and asked to provide a best estimate of size based on a series of intervals. Sixty-six percent of the organizations were coded as having the smallest membership level (below 100 members); we coded the remaining organizations as "1." Control of territory was defined as an organization being able to coerce non-member civilians to act or forebear, and to exclude police and military units from some defined geographic space over a period of time greater than six months. Eleven percent of the 395 organizations were found to have exerted strong control of territory at some point during the period of the study.

To represent network connectivity (as motivated in Section 2.2), we used data on alliances among the 395 organizations (see Asal et al., 2012: 40). We code network degree (**D**) as zero if an organization is recorded as having no ties to another group (which was the case for 46% of the 395 organizations), otherwise as a "1."

To represent knowledge and skills, we were able to use information from the BAAD-1 database as to whether each group was known to have either used or pursued chemical, biological, radiological, or nuclear (CBRN) weapons during the 1998–2005 period. This was the case for 3.5% of the 395 organizations.

Detailed information on each organization's professed ideology is available in the BAAD-1 database. To code the presence of ethnic ideology (**E**) we included organizations in which ethnicity accounts for any element of their founding, but we excluded groups that were founded on religious ideology. Twenty-three percent of the groups in our analysis were founded on an ideology that had an ethnic, but non-religious, component.

Four of our variables are dichotomous, and as described above we have dichotomized two others (size and network centrality measured by degree, **S** and **D**). We did so in order to maintain comparability with the "crisp set" analysis used early on by Ragin as a basis for his thinking (though he has subsequently generalized his approach to fuzzy sets; Ragin, 2000). Although the dichotomies make our analyses simpler to conduct and to convey to the reader, we are well aware that dichotomizing variables limits the generalizability of our findings. For example, if there are nonlinearities such that only very large or very small organizations involved themselves in the drug trade, our findings would clearly be distorted owing to the dichotomization. As mentioned previously (Section 1.1), this paper should not be read as a definitive study of terrorist groups and drugs, but rather as an effort to demonstrate the potential usefulness of our approach for generating insight on this subject.

With five factors included in our analysis—size (**S**), control of territory (**T**), network degreee (**D**), use or pursuit of CBRN weapons (**C**), and ethnic ideology (**E**)—there are 2^5 (32) configurations of these variables logically possible. We now study this space of configurations.

3.2. Ragin's QCA

In recent years, especially in the disciplines of sociology, political science, and related fields, there has been increased interest in an analytic framework developed by Charles Ragin and various research collaborators that goes under the name of configurational comparative methods (Rihoux and Ragin, 2009) or Qualitative Comparative Analysis (QCA) techniques (e.g., Ragin, 2000, 2008; Byrne and Ragin, 2009). Configurational analysis has its origin in efforts to bridge the gap between qualitative (typically case study oriented) and quantitative (variable oriented) approaches. This goal is achieved by utilizing Boolean algebra to study complex conjunctions of causal conditions across the cases in a study, working with a set-theoretic logic (for both crisp and fuzzy sets) and procedures that endeavor to reduce the complexity by identifying a relatively small set of configurations of causal conditions that, taken together (in their set-theoretic union), apply maximally across all the cases. This allows analysts to retain the relatively rich degree of complexity associated with in-depth, qualitative research while simultaneously obtaining relatively parsimonious explanations by evaluating the multiple conjunctions of factors that contribute to an outcome across multiple cases.

Grofman and Schneider (2009), in the course of providing an excellent introduction to QCA methodology, highlight several central contributions of configurational analysis that define it as a unique departure from standard statistical approaches. Focusing on "crisp set" analysis (where all variables are binary-coded), Grofman and Schneider outline several key features of QCA.

As mentioned briefly above, perhaps one of the most central elements of configurational analysis is the initial expectation that logical conjunctions of conditions—not single variables in isolation or in additive combination—are causally relevant for producing the outcome of interest. The emphasis of configurational analysis is on explaining cases, and there is no expectation that the same (appropriately weighted) set of factors will explain all the cases. This emphasis on equifinality (the notion that different combinations of variables might be associated with the same outcome) goes hand in hand with the assumption of multifinality. Multifinality presumes that the same variable can play different roles in different contexts. For example, a QCA equation of the form $CP + Up \rightarrow W$ might imply that countries exhibiting a generous welfare state (W) might have the *presence* of a leftist party (P) in combination with a corporatist polity (C), or the *absence* of a leftist party (p) in combination with strong unions (U). Some cases resulting in W would be characterized by CP while others would exhibit Up (as illustrated by Grofman and Schneider, 2009).

A final point that is key to our application of this approach is that configurational analysis allows analysts to move directly between individual cases and solutions derived from the aggregate data. As Grofman and Schneider (2009: 669) note, while regression modeling is oriented to finding the best overall fit across all the cases (according to well known least-squares criteria), configurationl analysis endeavors to find the closest to a 100% fit for as many cases as possible. In this way, each individual case accounted for by a solution will contain all the factors associated with that

[1] As of 2003, the US Drug Enforcement Administration ascertained that 14 of the State Department's then-current list of 36 designated foreign terrorist organizations had some degree of connection with drug activities (Casteel, 2003). Our dataset contains over twice as many groups (35) engaging in drug production or trafficking activities. Our much larger sample (395 organizations) includes many groups of lesser overall size and scope on average than those designated by the State Department in 2003.

solution. While configurational analysis is a deductive approach with models specified based on existing theory, this ability to move directly between solutions and cases facilitates continued inductive research, as groups of cases are co-classified based on a set of common factors. This allows researchers to return to the cases following analysis to better understand the implications of the solutions provided.

3.3. Configurational correspondence analysis: a two-mode network approach

Configurations of variables may be defined by the cases that exhibit them, even as cases may be understood as the factors or variables that comprise and define them. It is this "duality of cases and variables" that we focus on in conceptualizing configurational analysis as a form of two-mode network analysis (see also Butts, 2009 on the advantages of expanding analytical thinking and modeling repertoires as to what constitutes a "node" and what is an "edge" in network analysis).

3.3.1. Disjunctive coding

The number of variables in the analysis is doubled, by including each property as well as its complement. If property X has values in the range of 0–1, then a new property, defined as the complement of X and denoted by a lower-case letter, x, is defined as $x = 1 - X$ and is adjoined as a new column to the cases-by-properties data table. Thus disjunctive coding is appropriate for both crisp-set and fuzzy-set variables, though for simplicity we consider only crisp-set variables in this paper. This disjunctive coding is consistent with Ragin's insistence that modeling the absence of a property (such as being single) is not automatically symmetric with modeling the property's presence (such as being married).

3.3.2. Moving from two-mode to one-mode analysis: Ragin's measures of consistency and coverage

Let us denote by the symbol $\mathbf{X}$ the data table (395 cases by 12 properties, including each of the six binary variables and its complement). The matrix product $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, where $\mathbf{X}^{\mathrm{T}}$ is the transpose of $\mathbf{X}$, produces a square matrix showing overlaps among the variables (see Table 1). This operation has long been employed in the study of affiliation networks (Breiger, 1974). In the present context, we can use the result to see how two of Ragin's most important measures relate to two-mode network analysis.

Let us consider as an "outcome" the participation of our organizations in the drug trade (our variable $\mathbf{Y}$). Let us further consider a single predictor variable such as pursuit of CBRN weapons (variable $\mathbf{C}$). For these crisp-set (binary) variables, Ragin (2008: 50–54) defines his measure of set-theoretic *consistency* as "the sum of the *consistent* membership scores" of $\mathbf{C}$ within $\mathbf{Y}$, in ratio to "the sum of all the membership scores" of $\mathbf{C}$. From the overlaps depicted in Table 1, notice from cell [2,1] that there are 9 organizations coded both as pursuing CBRN (variable $\mathbf{C}$) and as engaging in the drug trade (variable $\mathbf{Y}$). From cell [2,2] we see that the total number of organizations coded as pursuing CBRN is 23. The ratio, 9/23 = .39, is the descriptive conditional probability of engaging in the drug trade, given that the organization is coded as pursuing CBRN weapons. This is precisely Ragin's measure of set-theoretic consistency, and in the case of single-variable predictors we can compute all possible consistency scores from Table 1. It is found, for example, that 26 organizations coded as *not* pursuing CBRN weapons participate in the drug trade (cell [8,1]), out of 372 organizations coded as not pursuing CBRN (cell [8,8]). Therefore the consistency score for not pursuing such weapons in predicting $\mathbf{Y}$ is 26/372 = .07.[2] We learn that, using Ragin's descriptive measure of consistency, groups pursuing CBRN are more likely than groups not pursuing such weapons to engage in the drug trade. (These are single-variable relationships. Ragin is particularly interested in configurations of variables, a topic to which we will soon devote attention.)

The second measure of primary importance in Ragin's QCA is *coverage*. How well does a variable (or a configuration of variables) do in accounting for all cases of the outcome? Continuing with the example of CBRN use or pursuit, we have seen that 9 organizations coded as using or pursuing CBRN weapons also participate in the drug trade (cell [2,1] of Table 1). The total number of organizations in the drug trade in this sample is 35 (see cell [1,1]). The ratio, 9/35 = .26, is the *coverage* of the set of outcomes ($\mathbf{Y}$) by the single predictor variable $\mathbf{C}$, illustrating Ragin's (2008:54–63) measure of set-theoretic coverage in the case of crisp-set analysis.

3.3.3. Barycentric correspondence analysis

The procedure described above works well for exploring the effects of single variables on the outcome. However, the comparative advantage of Ragin's QCA comes from discovering smallest sets of combinations of variables (each combination of variables is termed a "configuration" by Ragin) that have highest consistency while best covering the outcome. Ragin (2000: 120–145) describes a procedure that involves examining all combinations of variables. For our data, the 5 predictor variables generate $3^5 - 1 = 242$ combinations (Ragin, 2000: 127) that would need to be examined with respect to consistency (as defined above: the conditional probability of yielding the outcome, given the configuration of predictors). We seek a shortcut: a procedure less principled than examining all possible combinations of variables, and one that exploits the mutual constitution of cases and variables.

Because set-theoretic relations among the properties are at the center of our concern, a variety of techniques might be considered. Schweizer (1996) used both lattice analysis and a form of QCA on cases-by-property data. A main advantage of this Galois or "dual" lattice analysis (see also Mohr and Duquenne, 1997; Mische and Pattison, 2000) is that intersections and unions among the "cases" are produced in addition to orderings among the "properties" and, moreover, the two sets of intersections and unions are dual in the sense that orderings internal to each mode (cases and properties) are contained within a single lattice (see also Martin, 2006; Schaefer, 2010). Algebraic techniques for "factoring" such a lattice can then be used to reduce its complexity (e.g., Mische and Pattison, 2000). Alternatively, the analyst might consider a dimensional reduction of the cases-by-properties table, perhaps using correspondence analysis (e.g., de Nooy, 2003) to preserve the duality by representing cases and properties within the same space. In fact there are important relations between algebraic and dimensional reductions of a cases-by-properties table (Pattison and Breiger, 2002).

Within the community of social network researchers, correspondence analysis has provided a productive set of techniques for analyzing two-mode affiliation networks and bipartite graphs (Faust, 2005). Breiger (2009) recommends a method of analyzing the cases-by-properties table that uses a form of correspondence analysis that preserves important set-theoretic properties that are key to QCA as well as to the lattice analyses just mentioned. Known as barycentric correspondence analysis (Greenacre, 1984; Le Roux and Rouanet, 2004), this method locates each case in a reduced-form dimensional space at the intersection (or "barycenter," a term

[2] The two-mode formulation extends naturally in the case of fuzzy sets (when we consider single predictors). The matrix multiplication operators of multiplication and addition are replaced (respectively) with those of *min* and *sum*.

Table 1
One-mode matrix ($\mathbf{X}^T\mathbf{X}$) derived from two-mode data matrix (395 groups by 12 properties).

	DRUG	CBRN	DEGR	TERS	ETHN	SIZE	drug	cbrn	degr	ters	ethn	size
DRUG	**35**	9	27	14	18	29	0	26	8	21	17	6
CBRN	9	**23**	20	7	4	20	14	0	3	16	19	3
DEGR	27	20	**212**	30	46	92	185	192	0	182	166	120
TERS	14	7	30	**43**	12	39	29	36	13	0	31	4
ETHN	18	4	46	12	**91**	43	73	87	45	79	0	48
SIZE	29	20	92	39	43	**134**	105	114	42	95	91	0
drug	0	14	185	29	73	105	**360**	346	175	331	287	255
cbrn	26	0	192	36	87	114	346	**372**	180	336	285	258
degr	8	3	0	13	45	42	175	180	**183**	170	138	141
ters	21	16	182	0	79	95	331	336	170	**352**	273	257
ethn	17	19	166	31	0	91	287	285	138	273	**304**	213
size	6	3	120	4	48	0	255	258	141	257	213	**261**

Table 2
Barycentric scores for variables.

Variable	Dim. 1	Dim. 2
DRUG = **Y**	3.7920	−1.6955
CBRN = **C**	3.8054	3.4208
DEGR = D	0.6651	1.0377
TERS = **T**	3.2222	−0.1103
ETHN = **E**	0.7942	−3.5361
SIZE = **S**	1.8394	−0.1008
drug = **y**	−0.3687	0.1648
cbrn = **c**	−0.2353	−0.2115
degr = **d**	−0.7705	−1.2021
ters = **t**	−0.3936	0.0135
ethn = **e**	−0.2377	1.0585
size = **s**	−0.9444	0.0518

referring to the center of gravity) of the properties manifested by that case. If some specific case manifests properties A and B but not C (denoted ABc), then the dimensional location of that case (on each dimension considered separately) is simply the mean of the locations of properties A, B, and c.[3]

To illustrate the set-theoretic and dual scaling features of barycentric correspondence analysis, we provide the scores for all of our variables (Table 2) and for an illustrative subset of our 395 organizations (Table 3). Consider the first organization listed, which is the Liberation Tigers of Tamil Eelam (LTTE). As shown in the "Characteristics" column, the Tamil Tigers are coded in our database as participating in the drug trade (variable **Y**), exhibiting strong control of territory (**T**), having a founding ideology that emphasized ethnic but not religious solidarity (**E**), having large size (**S**), but not involved in CBRN activities or having high network degree (lower-case variables **c** and **d**).

Consider the score estimated for the Tamil Tigers on the first dimension of the barycentric correspondence analysis solution: 1.4403 (shown in Table 3). The crucial point is that this score is identical to the mean of the properties that the group possesses. Specifically:

$$\begin{aligned}&\text{Mean}\,(\mathbf{Y}, \mathbf{c}, \mathbf{d}, \mathbf{T}, \mathbf{E}, \mathbf{S}\ \text{on dimension 1})\\ &= \text{Mean}\,(3.7920, -0.2353, -0.7705, 3.2222, 0.7942, 1.8394)\\ &= 1.4403 \qquad (1)\end{aligned}$$

This result is general: barycentric correspondence analysis (CA) insures that each case is assigned a location within a multidimensional space at the mean of all its properties.

[3] Wolff and Gabler (1998) provide lattice analysis and barycentric correspondence analysis of identical 0–1 data in a cases-by-properties format, in a compelling example showing that the correspondence analysis provides a simplification of the full data that is consistent with the dual lattice of intersections of properties and cases.

The correspondence analysis model may be written as follows (Goodman, 1996: 412):

$$F_{ij} = Np_{i+}p_{+j}(1 + \lambda_{ij}) \qquad (2)$$

where F_{ij} are fitted cell values for the data table X (cases by variables), N is the table sum and p_{i+} and p_{+j} are the proportion of cases in each row and in each column. The λ_{ij} are the interaction terms relating cases to variables. Notice that if all $\lambda_{ij} = 0$ then Eq. (2) devolves to the simple independence model. Above and beyond independence, the CA model posits that the interaction structure of the table may be represented by sets of dimensions,

$$\lambda_{ij} = \sum_{m=1}^{M} r_{im}c_{jm} \qquad (3)$$

where the analyst posits M dimensions for rows (the r scores) and for columns (the c scores). Rewriting Eq. (2) to insert the λ_{ij}, we see that, in the spirit of the definition of the modularity matrix used for community detection in one-mode networks (Newman, 2006), the λ_{ij} provide measures of two-mode structure above and beyond that predicted by the independence model.

We are particularly interested in the interaction of each of our cases (terrorist organizations) with the outcome variable (**Y**, participation in the drug trade). From the r and c scores given in Table 2, we may compute λ_{iY}, the interaction of each case with **Y**. For the Tamil Tigers, for example, their interaction with the drug trade is, applying Eq. (3), $\lambda_{iY} = (1.4403 \times 3.7920) + (-1.1427 \times -1.6955) = 7.3992$ as shown in the final column of Table 3.

Just as the off-diagonal entries in Table 1 report overlaps among variables, the interactions illustrated in the final column of Table 3 report associations between cases and the outcome variable. Moreover, as each case is reduced to the configuration of properties it manifests, these interaction terms also relate configurations of case properties to the outcome variable.

3.3.4. *Visualization of configurations and their relations*

By the duality shown in Eq. (1), cases are analytically defined by the properties that comprise them. Cases and configurations may therefore be represented in the same space. Moreover, any two configurations that are dissimilar from each other (in the sense that each configuration tends to "cover" or include different sets of cases) will tend to be on opposite "sides" of the space, in a manner that will be illustrated in our data analysis. Thus, where QCA produces a list of configurations and information about them, the barycentric correspondence analysis emphasizes the relations among the configurations (and simultaneously the cases).

These two approaches to configurational analysis, QCA and barycentric correspondence analysis, complement each other. The tendency of QCA is to complicate; as Ragin (2008) points out, each new causal condition doubles the size of the analytic space formed

Table 3
Barycentric scores for selected groups.

Group	Characteristics	Scores Dim. 1	Dim. 2	λ_{iY}= Interaction with **Y**
Tamil Tigers (LTTE)	YcdTES	1.4403	−1.1427	7.3992
Jemaah Islamiya (JI)	YCdTES	2.1811	0.6017	7.2503
Abu Sayyaf (ASG)	YcDTeS	1.5076	−0.0037	5.7231
Irish Nat'l Liberation Army	YcDtEs	0.6130	−0.7234	3.5510

by these conditions. Ultimately, there must be some redundancy among the selected causal conditions, and correspondence analysis, as a data reduction technique, exploits that redundancy in a way that the more principled QCA cannot—by design. The barycentric correspondence analysis sets up a procedure for seeing one's way to the selection of cases for further in-depth study.

4. Results

Using the barycentric scores shown in Table 2 for the variables and illustrated for four cases in Table 3, we constructed a joint space for situating all 395 cases and all variables (Fig. 1). Table 2 tells us that the first dimension presents a contrast between, on the left side of Fig. 1, groups with small size and/or low network degree centrality (scores of −.9444 and −.7705) and, on the opposite side, groups with high propensity to use or pursue CBRN weapons, engage in the drug trade, and/or exert strong control of territory (scores of 3.8054, 3.7920, 3.2222). The strongest contrast defining the second dimension of Fig. 1 is that between ethnic (but non-religious) founding ideology (score of −3.3561) and CBRN activity (3.4208).

4.1. *Configurational correspondence analysis results*

We find that 22 of the 32 possible distinctive configurations of predictor variables are observed in this data. These 22 configurations are shown in Fig. 1 according to whether each configuration of variables contained any drug groups (denoted "+") or not (denoted "*").

Moreover, by constructing this "dual space," each organization is located at the center of gravity (or barycenter) of the variables it manifests.

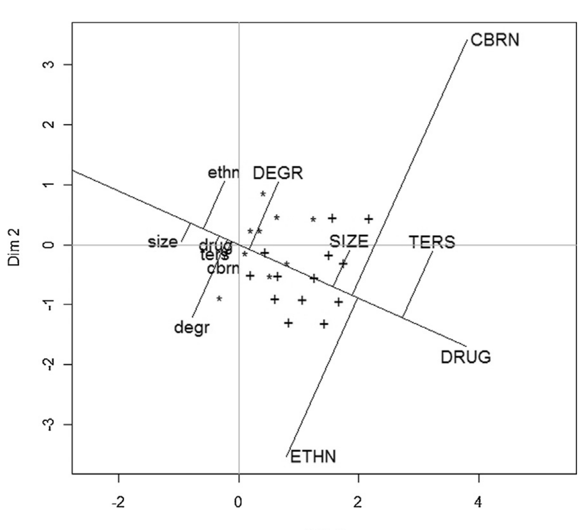

Fig. 1. The space of cases and configurations. Configurations containing drug groups are denoted "+".

We seek to identify specific configurations of variables that are most highly associated with an organization's engagement in the drug trade. In interpreting the structure of this space, a crucial result (Goodman, 1996; reviewed in Breiger, 2009) is that projections of points to a line connecting "DRUG" (variable **Y**) to the origin are proportional to the interaction terms λ_{iY} discussed and illustrated above. Therefore, all variables to the right of the origin are positively associated with participation in the drug trade, and the association increases as we follow projections of each variable to the line (which is shown in Fig. 1) that connects the origin to the outcome variable. Relatively high network centrality (degree centrality) tends to be a "lowest common denominator" for groups engaging in the drug trade, in that this variable is positively but not strongly associated with participation.

Fig. 1 suggests two general causal paths, in that ethnicity (**E**) is on the opposite side of the space from the other variables. Beginning at the origin and moving toward the dependent variable ("DRUG" for the presence of drug-smuggling), there is a suggestion that network connectivity ("DEGREE" or **D**) is pervasive among drug-smuggling groups, followed by size ("SIZE" or **S**) and then the other variables. This patterning is further suggested by Fig. 2, which reports the same "space" as the previous diagram but which identifies particular configurations within which there is a high probability that an organization engages in drug smuggling. (As an example of reading the configuration labels in Fig. 2, "CDST" stands for the conjunction of the use or pursuit of CBRN weapons [**C**], network connectivity [**D**], large size [**S**], and strong control of territory [**T**].)

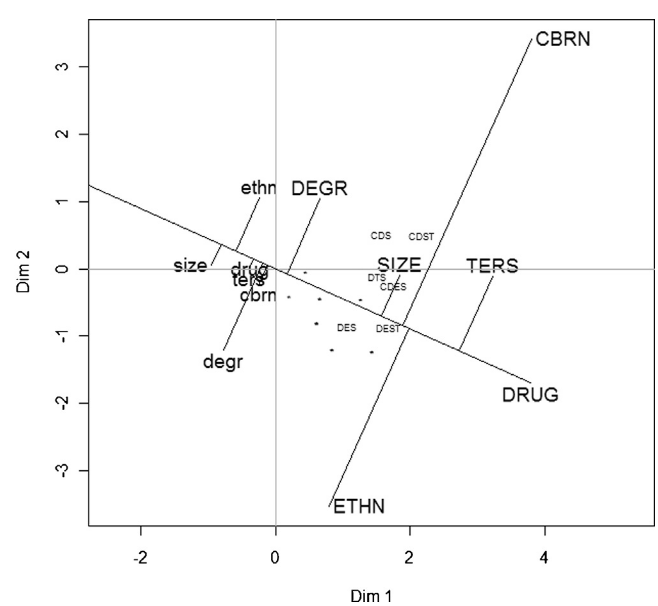

Fig. 2. Configurations strongly related to engagement in the drug trade. For the configurations shown, **D** = network centrality (degree); **S** = size; **E** = ethnic component to the group's ideology; **C** = CBRN use or pursuit; **T** = strong control of territory. For example, CDS means the intersection of **C** and **D** and **S**.

Thus, a particular set of configurations is suggested as predictive of drug smuggling by reading the variables from the origin toward "DRUG" in Figs. 1 and 2:

$$\mathbf{D}(\mathbf{S}+\mathbf{C}+\mathbf{E}+\mathbf{T})$$

$$\mathbf{DS}(\mathbf{C}+\mathbf{E}+\mathbf{T})$$

$$\mathbf{DSC}(\mathbf{E}+\mathbf{T})$$

where the five predictor variables are abbreviated as above, and where multiplication indicates set-theoretic intersection while addition indicates the (non-exclusive) union of sets, as in Ragin's approach (Ragin, 2000, 2008; Grofman and Schneider, 2009). The first configuration listed above consists of all two-variable conjunctions that include network connections (**D** for network degree greater than zero). The second consists of all three-variable conjunctions that include **DS** (network connections and large size, **DS**). The third consists of all four-variable conjunctions that include **DSC** (network centrality and large size and use or pursuit of CBRN weapons).

How shall we evaluate these configurational models (or others that we might wish to consider)? We begin with a baseline. If we know nothing except the dependent variable, we would find that 35 of our 395 groups engage in drug smuggling. The ratio 35/395 is 0.089, so there would be a 9% chance of correctly predicting that a group engages in drug smuggling if we were to select a group at random from these 395.

Consider now the first configuration listed above. Of our groups, 118 have either DS or DC or DE or DT (that is: network connectivity conjoined with one of the other predictor variables). Of these 118, 26 groups also engage in drug smuggling. Thus, if an analyst were to predict drug smuggling by singling out groups that manifested **DS + DC + DE + DT = D** (**S + C + E + T**), the analyst would be correct 26/118 = 22% of the time, up from the baseline of 9%. We think of 22% in this example as the conditional probability of correctly predicting drug-smuggling given the configuration **DS + DC + DE + DT**. As discussed above, Ragin (2008) calls this conditional probability "consistency" in the context of crisp set QCA.

Consider next the second configuration listed above. Of our groups, 53 have either **DSC** or **DSE** or **DST** (that is: network connectivity and also large size, both conjoined with at least one of either CBRN activity, an ethnic component to their ideology, or strong control of territory). Of these 53 groups, 22 of them also exhibit drug smuggling. Thus, if an analyst were to predict drug smuggling by singling out groups manifesting **DSC + DSE + DST = DS** (**C + E + T**), the analyst would be correct 22/53 = 41.5% of the time, up very considerably from the baseline of 9%.

The third and final configuration listed above is even more restrictive. Of our 395 groups, only 9 of them have network connectivity and large size and CBRN activity and *either* of ethnic ideology *or* strong control of territory, that is: **DSC**(**E + T**) = **DSCE + DSCT**. However, of these 9 groups, 7 of them also engage in drug-smuggling, for a conditional probability of 7/9 = 78%. Therefore, while an analyst predicting drug smuggling on the basis of this configuration would have a very high "hit" rate (78%), only 20% of the drug groups (7/35) would be included in the set of groups correctly predicted.

As discussed above (Section 3.3), Ragin (2008) terms the latter fraction the "coverage ratio." Coverage scores for the three configurations discussed above are 26/35 (=74%), 22/35 (=63%), and 7/35 (=20%), respectively.

Looking jointly at the conditional probabilities of correct prediction (the "consistency ratios") and at the coverage ratios, we are drawn to the second configuration discussed above:

$$\mathbf{DS}(\mathbf{C}+\mathbf{E}+\mathbf{T}) = \mathbf{DSC}+\mathbf{DSE}+\mathbf{DST}$$

This configuration has a consistency ratio of 41.5%, which is 4.68 times the baseline prediction of 9%. At the same time, this configuration covers 63% of all the cases exhibiting drug smuggling.

This analysis well illustrates equifinality, or the notion that multiple configurations might lead to the same outcome. Our best prediction of drug smuggling is for groups that have network connections (**D**) and noticeable size (**S**), in combination with at least one of CBRN activities (**C**), an ethnic component to the group's ideology (**E**), and strong control of territory (**T**). In Section 5, we devote considerable attention to the interpretation of these findings.

4.2. Confidence intervals for consistency and coverage measures

Our major descriptive finding is that, for the configuration DS (**C + E + T**) = **DSC + DSE + DST**, the conditional probability of correctly predicting engagement in the drug trade is 4.68 times (=[22/53]/[35/395]) the baseline. Let us call 4.68 our observed Improvement Ratio.

We would like to place a confidence interval around our descriptive finding. To convey our approach for doing so, let us consider by analogy the problem of placing a confidence interval around the estimate of a simple regression coefficient computed from a random sample. We could imagine taking a second random sample from the same population, and computing a regression coefficient from this second sample. We could in fact imagine repeating this procedure many times, thus obtaining a long list of estimated regression coefficients in this manner. We could consider as the 95% confidence interval the interval within which the central 95% of those estimated regression coefficients were located. Usually, we go to much less trouble than this because, if the assumptions of regression analysis are met (crucially including the assumption that errors of prediction have a Gaussian distribution), we can use an explicit formula to calculate the standard error and then the confidence interval. However, if the usual assumptions are not met, but if we are still willing to assume that our data are sampled randomly from the population of interest, we can nonetheless estimate a 95% confidence interval by using the bootstrap method due to Efron that has been reviewed in this paragraph (Efron, 1982). As Diaconis and Efron (1983) demonstrate, the bootstrap offers an effective alternative to having to meet the assumption that data conform to a Gaussian distribution, as well as to having to rely on statistical measures that can be formulated only as closed-form equations. We are not aware of previously published bootstrap estimates of confidence intervals for key quantities computed from QCA or from other configurational approaches, but we consider the construction of such estimates to be a contribution toward assessing the effect of sampling variability on our key finding.

We are willing to make the assumption that our sample of 395 terrorist organizations is a random sample from a population of interest: all terrorist organizations that we would find if we hired many different teams to apply our coding rules to data that they would collect using our case selection procedures. Therefore, now considering our 395 groups as a population, we sampled 395 groups, with replacement, from our 395 groups, and we computed the Improvement Ratio for this bootstrapped sample. (That is, for this second sample of cases, we computed the baseline rate of involvement in the drug trade; we computed the descriptive conditional probability that cases within our favored configuration, **DSC + DSE + DST**, were involved in the drug trade, and we computed an Improvement Ratio that compared the conditional probability of

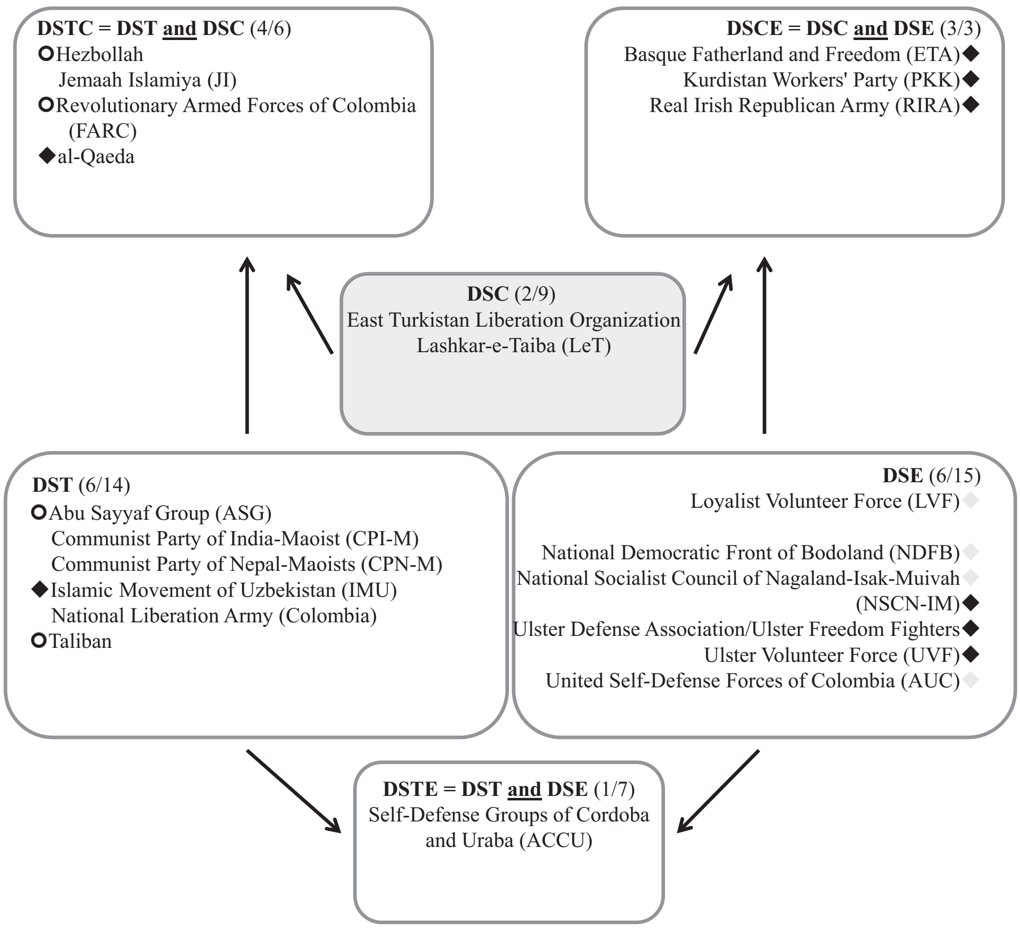

Fig. 3. Configurations of variables, and drug-active groups within each configuration (22 drug groups). **D** = network centrality (degree), **S** = size, **T** = strong control of territory, **C** = CBRN use/pursuit, **E** = ethnic ideology. (Fractions in parentheses = number of drug-active groups/all groups out of 395 having that configuration.) Arrows join unions of sets.

drug involvement to the baseline found for this second sample.) Of course the Improvement Ratio computed from this sample was not equal to the one we observed; however, we are interested in how much variation due to sampling variability is present. We therefore repeated this procedure 10,000 times. Based on this bootstrapped distribution of Improvement Ratios, we estimated a 95% confidence interval for our Improvement Ratio. The confidence interval indicated that, 95% of the time, the Improvement Ratio is between 3.56 and 6.21 times the baseline, which is a highly significant improvement. We are thus able to conclude that our observed Improvement Ratio (4.68) is significantly greater than its expectation (1.00) after the role of sampling variability is taken into account.[4]

Our observed coverage proportion (the percentage of all drug smuggling cases that we correctly predicted) was 63% (=22/35). The 95% confidence interval for our estimate of coverage (based on our 10,000 bootstrapped samples) is 46–79%.

Producing bounds for our key estimates also indicates the limits of our contribution. For example, the bounds on the Improvement Ratio translate into a conditional probability (consistency score) ranging from 31.5% to 55% that our favored configuration of variables will identify groups engaged in the drug trade. While high in comparison to the baseline (9%), and perhaps demonstrating a good starting point that can be improved upon in subsequent research, these results make clear the long road that still needs to be traveled in order to produce adequate indicators of drug trade engagement on the part of terrorist organizations.

5. Interpretation: using these results for inductive hypothesis generation

A good configurational analysis identifies anomalous cases and leads to new questions that impel the analyst to formulate a new configurational model, in an iterative process that, if successful, converges on improved understanding (Ragin, 2000, 2008).

Fig. 3 details the cases in each of the configurations that our analysis emphasizes: **DST**, **DSC**, and **DSE**, as well as set-theoretic unions of these three configurations. The 22 organizations implicated in these configurations are named in Fig. 3. (Twelve of these groups are marked with open circles or diamonds, and these symbols will be explained presently.) The duality inherent in our approach (treating cases as being constituted by the variables they exhibit, just as the configurations are constituted by the cases they represent) allows us to develop hypotheses based on commonalities that may have been previously unseen.

To demonstrate this process, we draw directly on the work of Felbab-Brown's (2010) Brookings Institution study. In her Appendix A, Felbab-Brown provides a table that lists organizations that either are, or have been involved in the drug economy, and she annotates the nature of their involvement. There are numerous

[4] Other sources of variability should be taken into account, for example that larger organizations are more likely to find their way into our sample. We have also explored bootstrapped samples where the probability of appearing in the sample is proportional to the size of the organization. Results are available upon request from the first-listed author. See Appendix A for further discussion.

Table 4
Comparison of cases representing configurations with drug activity presented by Felbab-Brown (2010).[a]

Configurational solution	Type of engagement in drug trade
DSCT	
al-Qaeda	Trafficking
Hezbollah	Trafficking; taxation of cultivation and distribution
FARC	Trafficking; taxation of cultivation; and processing
DST	
Abu Sayyaf Group (ASG)	Production
Islamic Movement of Uzbekistan (IMU)	Trafficking
Taliban	Trafficking; processing; taxation of cultivation
DSCE	
Basque Fatherland and Freedom (ETA)	Trafficking; distribution; money laundering
Real Irish Republican Army (RIRA)	Trafficking; distribution
Kurdistan Workers' Party (PKK)	Trafficking; taxation of trafficking; money laundering; and processing
DSE	
National Socialist Council of Nagaland-Isak-Muivah	Trafficking
Ulster Defence Association/Ulster Freedom Fighters	Trafficking; distribution
Ulster Volunteer Force (UVF)	Trafficking; distribution

[a] Only cases found in both our analysis and Felbab-Brown's (2010) analysis were included in the construction of this table.

reasons why her list and ours are not entirely comparable.[5] With this qualification in mind, Table 4 nonetheless lists, for all groups where there is correspondence between the two datasets, the location of each organization within the configurational space of our approach (as reported in Fig. 3) and summarizes Felbab-Brown's data on the nature of their involvement.

We have mapped Table 4 onto Fig. 3 in the following way. The twelve cases in Fig. 3 that are marked either by an open circle (○) or by a diamond (⧫) are in Felbab-Brown's list (Table 4). The four organizations involved in drug cultivation or taxation of cultivation are marked by open circles.[6] The others, all of which are involved in trafficking, are marked by diamonds.

All organizations listed in Fig. 3 have high network degree (**D**), large size (**S**), and at least one other of our five attributes, and they all engage in drug activities. Those listed on the left side of Fig. 3 are coded as exhibiting strong control of territory (**T**), while those on the right have a secular (not specifically religious) ideology emphasizing ethnic elements (**E**). Notice that, the left side of Fig. 3, which lists large and well-connected groups that control territory (**T**), four of the six groups listed by Felbab-Brown have drug activities involving cultivation or its taxation (open circles, as opposed to diamonds). Notice further that, on the right side of the same figure, listing large and well-connected groups that manifest an ethnic ideology (**E**), zero of the six organizations listed by Felbab-Brown engage in cultivation or its taxation, though all are involved in trafficking.

The implication that strong control of territory is necessary for a terrorist organization that cultivates drugs or taxes cultivation is not startling, at least at first glance. However, according to Felbab-Brown's (2010: 21) political capital theory of illicit economy, these groups—the ones that grow drug crops or levy taxes on cultivation—are the ones able to derive the highest degree of political capital from the drug trade, owing to their ability to protect the local economy and win the confidence of the surrounding populace. They are the ones at which the government policy of eradication is most often aimed, even though Felbab-Brown argues that such a policy is typically unsuccessful (because it strengthens the bond between cultivators, the local population, and the belligerent groups that protect them, among other reasons). We learn from our interpretation of Fig. 3 that many other groups—most notably, those whose ideologies emphasize a strong ethnic component—are engaged in aspects of the drug trade that do not exploit cultivation directly. We hypothesize that the ethnic ideology of these groups leads to trust (as argued in the social capital theory of Coleman, 1988, as well as in the social identity theory of Tajfel, 1981), and that trust is a strong enabler of covert trafficking in illegal commodities.[7]

In brief, we have found not one, but multiple "recipes" for terrorist groups engaged in drug activities. Being a large and well-connected group is highly important in most cases. Beyond that, some drug-active groups control territory, tend to exploit directly the process of cultivation, and are often targets of (unsuccessful) government policies aimed at eradication. Other drug-active terrorist groups do not control territory but have a strong ethnic, non-religious component to their ideologies, of the sort that we hypothesize encourages trust; these groups engage in drug trafficking and distribution. In brief: our results suggest that there is one type of drug-active terrorist group that follows a logic of territorial control, and a distinct type that seeks to discover ephemeral resources, exploit them quickly, and then disperse to search for new resources. A similar emphasis on the roles of two distinctive logics arose in the population ecology study of ethnic group boundaries of Lauwagie (1979).

In a departure from the argument of Felbab-Brown (2010), whose work we consider to be path-breaking and highly insightful, we observe that those terrorist groups that traffic in drugs but do *not* control territory and that are in many cases far removed from the drug growing fields may in fact be susceptible to attack by a policy of eradication. This follows from our friendly critique of Felbab-Brown that uses a corollary of her main thesis, which is that, by virtue of being removed from the territory of drug cultivation, the profits of these groups would decline as a result of reduced production levels but the groups would be unable to exploit the suffering of the local populations (in other countries) targeted by the eradication policy.

6. Conclusion

In this paper we have begun the generalization of an approach to two-mode network analysis so that it applies to the cases-by-variables data format. We have presented an approach that facilitates both descriptive research and the generation of hypotheses. By using barycentric Correspondence Analysis as a form of two-mode network analysis in manner inspired by Ragin's QCA, we

[5] Among the reasons for lack of complete comparability are: (a) Felbab-Brown has done deep and extensive fieldwork and has vast field experience, whereas we rely on a database; (b) her list covers two dozen countries whereas ours includes 65; (c) she only lists groups that engage in the drug trade, whereas 91% of the groups in our database do *not* engage in drug-related activities; (d) the two lists of groups are far from overlapping, with her list but not ours including "British government and traders, Colonial era," and others not qualifying as organizations in the database we draw upon, which is further restricted to the 1998–2005 period.

[6] We are uncertain about the classification of Abu Sayyaf (ASG). For example, a US Library of Congress report (Berry et al., 2002) states both that the group's involvement is "limited to sale and production" but also that "Philippine intelligence indicates that the ASG grows much of its marijuana on [two] islands" (Berry et al., 2002: 100, 105).

[7] Furthermore, in addition to finding that all six of the groups that Felbab-Brown lists as engaged in the drug trade and that we code as having important ethnic components to their ideology (marked by diamonds on the right-hand side of Fig. 3) are (according to Table 4) engaged in trafficking, we find that four of the six are engaged in distribution, which is a related illegal activity that benefits from having trusted associates. (Only one of the six groups *not* having an ethnic, non-religious ideological component is listed as engaged in drug distribution.)

hope to have cleared a path for future research that seeks to elaborate our understanding of how such networks can be used to model the occurrence of a specified outcome. Our approach provides visual simplification while retaining the conceptual complexities inherent in configurational research and modeling of complex statistical interactions (Brambor et al., 2006; Melamed et al., 2013). This retention of complexity preserves a fundamental element which underpins two-mode network analysis: the co-constitution of entities at different levels of structure (Breiger, 1974).

Our analysis found multiple configurations of variables that are associated with terrorist group participation in drug activity, rather than a single pathway. There are some common features that distinguish terrorist organizations engaged in the drug trade from those that do not (these common features being social-network connectivity with other terrorist groups and relatively large size). At the same time, however, the activity "signatures" of drug activity differ depending on whether these common features combine with strong control of territory, ethnic grievances, or pursuit or use of unconventional weapons. In this way we used our study of drug activity to illustrate one promising tool for more rigorous analysis of an issue of growing importance. We demonstrate how this conceptualization can be used not only to evaluate existing theories, but also to develop new hypotheses through the inductive re-analysis of configurations. Our findings point to future directions of inquiry but do not offer definitive directions for policy.

The ability to move directly between the result of this approach and individual cases distinguishes this approach from other methodological techniques that similarly aim to identify multiple trajectories. Semi-parametric group-based trajectory analysis (LaFree et al., 2010; Nagin, 1999) similarly facilitates the identification of multiple trajectories toward a specific outcome, but differs insofar as the method is semi-parametric. Trajectory analysis uses maximum likelihood estimation to identify clusters of cases that exhibit different trends within an independent variable matrix. Because this method approximates a continuous distribution using a discrete mixture of indicators, each trajectory only represents an approximation (Nagin, 1999). Our non-parametric approach to identifying equifinality allows us to move directly between the results of our analysis and the individual cases that each solution represents, thereby retaining the individual significance of cases within the context of our analysis and facilitating the inductive re-analysis discussed above. As such, our approach represents a necessary bridge between the generalized approximations produced by models that rely on net effects, and the deep knowledge associated with individual case studies. Because we build on Ragin's (2008) QCA approach, we are able to analyze multiple 3- and 4-way interactions that almost always elude standard regression modeling and that would be infeasible to specify within a regression context applied to the dataset and substantive concerns that we have pursued.

In addition to gains, there are analytic losses associated with our current formulation of the barycentric approach. We have no parallel to the net-effects modeling of regression analysis and its many generalizations, and we have much more limited ways of dealing with inferential issues (though we think our introduction of bootstrapping for configurational models, Section 4.2, is worthy of further work). It seems possible to combine the two-mode approach of this paper with regression modeling, as suggested only implicitly by Breiger et al. (2011), and that direction would be a desirable one to pursue. Furthermore, as elaborated in Section 3.1, there is the limitation in this paper of having dichotomized variables that would better have been analyzed at ordinal or interval levels, an issue we pursue further in Appendix A.

It is important to explicitly note therefore that as we leverage points of convergence between multiple methods, we do lose elements that each method independently brings to the table. However, whether the data represent events by actors, individuals by organizations, or any other two-mode relationship, the ability to systematically evaluate commonalities within the data in relation to a given outcome while simultaneously presenting a simplified visual map of the social space represented by the data should prove to be of great utility.

Acknowledgments

We are grateful for discussions with Gary Ackerman, Elisa Bienenstock, Moses Boudourides, Jennifer Earl, Brint Milward, Joerg Raab, Charles Ragin, John Tidd, and Joseph West. We also thank the special issue editors and referees for providing many constructive comments and suggestions. This work was supported by the Defense Threat Reduction Agency (Basic Research Award HDTRA1-10-1-0017) and the Air Force Office of Scientific Research (MURI Award FA9550-10-1-0569). Eric Schoon was supported by a National Science Foundation Graduate Research Fellowship.

Appendix A. Issues pertaining to variable coding, our size variable, and newspaper sources

We provide a focused treatment of several issues in this appendix: greater detail about the dataset and the coding of the variables we use; an assessment of the decision to binarize the size variable; the conformity of our data with the present state of the art in published terrorism research; and a statement of the potential problems of relying as heavily as we do on newspaper sources.

A.1. Collection of data and coding of variables

The BAAD-1 dataset ("Big, Allied, and Dangerous"), which is housed at and maintained by The National Consortium for the Study of Terrorism and Responses to Terrorism (START),[8] a U.S. Department of Homeland Security Center of Excellence, is currently the most extensive dataset on terrorist attributes and activities taking organizations as the units of analysis that is publically available. BAAD-1 synthesizes data from the Terrorism Knowledge Base (TKB; produced by the RAND Corporation for the Memorial Institute for the Prevention of Terrorism), the Correlates of War Project (Correlates of War, 2004; Singer, 1988), Polity IV dataset (Marshall et al., 2006), and the Monterey Terrorism Research and Education program at the Monterey Institute of International Studies Group Characteristics Database (Ackerman, 2006), in addition to providing new information based on original coding efforts (for further information, see Asal and Rethemeyer, 2008, 2009; Asal et al., 2009b).

All data in BAAD-1 was coded using the same general procedures. However, in some cases, extra validity checks were implemented due to variation in data availability. Below, we outline the coding procedures for each of the variables included in our analysis.

*A.1.1. Size (**S**)*

The variable for size was coded based on semantic data from the TKB. This variable was coded ordinally using the following coding scheme: 0–100 and low confidence = 0, 100–1000 = 1, 1000–10,000 = 2 and 10,000 or more = 3. While this ordinal scheme is less preferable than one featuring exact membership counts for each organization and may mask the effect of subtler size variations, the approach we review was adopted in an effort to retain a high degree of sensitivity to variation and missing data risks that

[8] http://www.start.umd.edu/start/data_collections/ (accessed 1/2/13).

are commonly associated with the study of terrorist organizations. Assigning groups to ordinal size categories assured the accuracy of the representation, which was of the highest priority in coding this data. Of the 395 groups in the dataset all except 77 had validated measures of size included in the TKB. For these 77 groups, all were expected to be small, but information that offered a precise size range was unavailable. Thus, the first category in the ordinal coding was listed as a size range from 0 to 100, or low confidence.

To validate the measure of size, the BAAD-1 researchers provided a list of all groups in the dataset to experts from the Monterey Institute for International Studies and members of the US Intelligence Community. These experts were then asked to provide a best estimate of the groups' sizes. This data was then cross-referenced with the data collected from the TKB to verify the coding of the variable. For cases where there was disagreement, the researchers looked for outside sources and went with the preponderance of the evidence (which underscores why the researchers used an ordinal scale rather than a specific number).

In Section A.2 we provide an assessment of our decision to binarize the Size variable in the analyses reported elsewhere in this paper.

*A.1.2. Network degree (**D**)*

Case histories included in the TKB included information on 22 types of affiliation that a given terrorist group might have. Using this data, coders identified the number of connections that each group in the dataset had to other known terrorist groups. Groups were coded as being allied to other groups if the TKB listed them as having positive ties to other groups (e.g., "allies" rather than "rivals"). The measure for connectivity used in this analysis is the degree centrality of each group based on the ties that were identified through this coding procedure.

In cases where inconsistencies were identified (e.g., one group having alliance with another group that did not have the same alliance listed), coders sought independent confirmation from available academic, intelligence or print media resources. Codes were assigned only once validity had been confirmed through these sources.

*A.1.3. Control of territory (**T**), ethnicity (**E**), and CBRN use and pursuit (**C**)*

As discussed in Section 3.1, above, control of territory was defined as an organization being able to coerce non-member civilians to act or forebear, and to exclude police and military units from some defined geographic space over a period of time greater than 6 months. The variable for ethnic ideology was coded as "1" if the organization is ethnonationalist but does not have a religious ideology. Finally, CBRN use and pursuit was coded as 1 if the group was known to have used or pursued chemical, biological, radiological or nuclear weapons. The coding and verification of these variables followed the same coding procedure. Each measure was drawn initially from the TKB profiles, and coding was verified using available academic, intelligence and print media resources.

*A.1.4. Drug trafficking (**D**)*

Groups were coded as being involved in drug trafficking if (a) the group was known to be involved in trafficking drugs, or (b) the group was identified as the original grower and supplier. Initial coding was based on the TKB data; however, because this was not a central variable in the TKB, further research was done on each group. Much of the data for this variable was drawn from news reports found through LexisNexis. LexisNexis provides access to news outlets worldwide, allowing our search procedure a higher degree of sensitivity. Because the groups in the dataset differed widely in prominence, there was wide variation in the degree of coverage for each group. However, before any case was coded as having smuggled drugs we sought verification from a minimum of two sources.

Coders were trained to look for explicit evidence of either trafficking or supplying. Here are the fictional examples used in the training text as examples that provide evidence of drug activity:

> "The KLA's involvement in drug smuggling as a means of raising funds for weapons is long-standing. Intelligence documents show it has aligned itself with an extensive organized crime network in Albania that smuggles heroin to buyers throughout Western Europe and the United States. Drug agents in five countries believe the cartel is one of the most powerful heroin smuggling organizations in the world."
>
> OR
>
> "France's Geopolitical Observatory of Drugs said the KLA was a key player in the rapidly expanding drugs-for-arms business and helped transport $2 billion in drugs a year into Western Europe."

The data was coded and validated using the same set of procedures as all other previously published data from BAAD-1 (see Table A.2 for a list of publications). Multiple sources to validate the coding of each case were required. All available information was sought for each case, using resources that ranged from academic books on drug smuggling and terrorism, to publically available government reports.

A.2. Assessing the decision to binarize the size variable

Section 3.1 states that we binarized the size variable in order to maintain comparability with the "crisp set" analysis that is still widely used in applications of Ragin's QCA methods in spite of his own concentration in recent work on the development of (generalized) fuzzy coding and on calibration schemes (Ragin, 2008). In Section 4.1 we present our preference for the configuration

$$\mathbf{DS}(\mathbf{C}+\mathbf{E}+\mathbf{T}) = \mathbf{DSC}+\mathbf{DSE}+\mathbf{DST}$$

We reported that the consistency score associated with this configuration was .42 (=22/53), and that the Improvement Ratio in comparison to the baseline probability that a group engaged in drug activities was 4.68 times the baseline ratio of .09 (=35/395). In Section 4.2 we placed a 95% confidence interval of (3.56, 6.21) around this estimate of the Improvement Ratio, finding it to be significantly different from 1.00 under the null hypothesis that the groups were sampled at random from a population of interest. We also found a coverage score of 63% (with a confidence interval of from 46% to 79%).

Now we raise the question of whether our results depended on our binarization of the size variable (**S**). Column A of Table A.1 restates the results reviewed just above. (The meaning of the final row will soon be explained.)

We ask: How would the performance of our preferred configuration (the one reviewed just above) change if we had used different codings of the Size variable. Ragin (2008) discusses moving from conventional quantitative variables to fuzzy sets by means of the process of calibration. Essentially, the analyst asks what values of the given variable are needed in order for a case to have a given level of membership (ranging from just above 0% to just under 100%) within the variable of interest. For example, in Column B of Table A.1 we consider that very small groups and those in which the analysts had low confidence as to knowing their size would have a very low membership in the size variable (which we operationalized as 5% membership in the size variable; see Chapters 4 and 5 of Ragin, 2008, for an extensive and informative discussion of calibration). Continuing with the coding in Column B, we

Table A.1
Assessment of the size variable under alternative codings.[a]

	Crisp or fuzzy coding				
	A	B	C	D	E
Category (number of members)					
0–100, and low-confidence	0.00	0.05	0.10	0.25	0.40
100–1000	1.00	0.20	0.40	0.50	0.60
1000–10,000	1.00	0.70	0.85	0.75	0.90
10,000+	1.00	0.95	0.95	0.95	0.95
Assessment					
Consistency	0.42	0.45	0.43	0.39	0.38
Improvement Ratio	4.68	5.05	4.83	4.37	4.24
Improvement Ratio CI lower bound	(3.56)	(3.69)	(3.65)	(3.39)	(3.31)
Improvement Ratio CI upper bound	(6.21)	(6.86)	(6.44)	(5.61)	(5.42)
Coverage	0.63	0.33	0.44	0.45	0.54
Coverage CI	(.46, .79)	(.19, .49)	(.27, .63)	(.28, .64)	(.34, .75)
Coverage Improvement Ratio	0.76	0.74	0.75	0.74	0.75

[a] 10,000 bootstrapped samples were run for each coding; resulting confidence intervals (CI) are 95%.

operationalized groups of size 100–1000 to have a moderate degree of membership (20%) in the size variable, whereas those with memberships estimated in the range of 1000–10,000 were considered to have a high degree of membership in size (70%), and groups of over 10,000 members even higher degree of membership in Size (95%).

Columns C through E provide alternative calibrations of the given four size categories with degrees of membership in the size variable. The alternatives in Columns B through E provide a very wide range of choices for calibrating this variable measured at the ordinal level, and they all contrast with the crisp-set coding in Column A, which is the one we used in this paper. Note in particular that Column D gives almost equal incremental weight to each additional size category. Column E considers even very small groups to have substantial degree of membership in the size variable.

What difference would these alternative codings make to our preference for our favored configuration (the one reviewed above)? With respect to the consistency scores, we find negligible difference when comparing the score of .42 which we reported in our main analysis (Section 4.1) with scores ranging from .38 to .45 for the alternative calibrations of the size variable (see Table A.1, comparing Column A with each of the remaining columns). We computed confidence intervals for the coding in each column of Table A.1 (each consisting of 10,000 bootstrap replications, as described in Section 4.2), and these too show essentially no difference between our binarization of the Size variable and the wide range of alternative, fuzzy calibrations as given in Table A.1.

The results in Table A.1 for coverage, however, do show that coverage decreases for each of the alternate codings. Compared to our report of a coverage score of 63% with our crisp-set coding (Column A), the alternatives exhibit coverage scores that are lower, ranging from 54% down to 33%. The 95% confidence intervals indicate that, in all but one instance, the declines in coverage are not severe. Moreover, recall that coverage is defined as the fraction that has in its denominator the number of groups in the sample (35 groups in all) coded as engaged in drug activities. However, if we consider weighting each organization by its degree of membership in the size variable (ranging across the interval from 0 to 1), we should perhaps adjust the denominator by weighting each of the organizations engaged in drug activities by its grade of membership in size. When we use this alternative denominator (see the concluding line of Table A.1) we find remarkably less variation in coverage scores with respect to alternative coding schemes for the size variable.

In brief: Table A.1 presents reasonable evidence to the effect that our binarization of the size variable did not affect the consistency score or the associated Improvement Ratio of the configuration of variables that we prefer in this paper (Section 4.1, reviewed above). However, there is also evidence that the coverage scores decreased under alternative codings, though perhaps not substantially so if we take account of the fact that different degrees of membership in the Size variable affect the denominator (number of groups engaged in drug activities, with such groups weighted by their size) of the coverage measure. Our measures reported in the final row of Table A.1 are, we think, well-taken, but they are not standard. Therefore the finding concerning coverage provides a further caution for us in interpreting the generalizability of our results.

A.3. Our dataset and the current state of the art

Writing in a symposium on political networks edited by Scott McClurg and Joseph K. Young, Perliger and Pedahzur (2011: 46) paint a canvass within which our dataset is well represented:

> "From a methodological perspective, contemporary students of political violence enjoy superior access to information and data about violent groups. Since September 11, growing numbers of media outlets have increased their coverage of terrorist incidents and groups. This rise in attention, combined with a striking increase in efforts and resources invested in data collection on these groups by academic and governmental agencies in recent years (e.g., START at the University of Maryland and TIGER at the University of Texas at Austin), have simplified the adaptation of research methods—such as SNA—that demand high-resolution information about terrorists and their groups."

Our BAAD-1 dataset, which is part of the START collection and the most extensive publicly available dataset that takes terrorist groups as the units of analysis, fits well with the current state of the art in the quantitative analysis of political violence and violent extremist groups. Research based on the BAAD-1 dataset has been published in a wide variety of journals including leading disciplinary journals (*Journal of Politics*, *Research in the Sociology of Organizations*, and the *Lecture Notes in Computer Science* series) and also leading journals in the specialty of terrorism and security studies (*Studies in Conflict and Terrorism*, *Conflict Management and Peace Science*). A list of research publications based on the BAAD-1 dataset is provided in Table A.2.

A.4. Issues in the reliance on newspaper sources

Extensive use of newspaper sources in the study of collective behavior and social movements has received a great deal of attention in recent times, with the review of Earl et al. (2004) often taken as definitive. An argument commonly made about newspapers is

Table A.2
Chronological list of BAAD-1 publications to date.

Asal, V, Blake, E.L., 2006. Creating simulations for political science education. Journal of Political Science Education 2 (1), 1–18.
Asal, V., Rethemeyer, R.K., 2006. Researching terrorist networks. Journal of Security Education 1 (4), 65–74.
Asal, V., Rethemeyer, R.K., 2008. Dilettantes, ideologues, and the weak: terrorists who don't kill. Conflict Management and Peace Science (3), 244–263.
Asal, V., Rethemeyer, R.K., 2008. The nature of the beast: terrorist organizational characteristics and organizational lethality. Journal of Politics 70 (2), 437–449.
Asal, V., Rethemeyer, R.K., Anderson, I., Rizzo, J., Rozea, M.M., Stein, A., 2009. The softest of targets: a study on terrorist target selection. Journal of Applied Security Research 4 (3), 258–278.
Asal, V.H., Rethemeyer, R.K., 2009. Islamist use and pursuit of CBRN terrorism. In: Ackerman, G., Tamsett, J. (Eds.), Jihadists and Weapons of Mass Destruction: A Growing Threat. Taylor & Francis, London.
Anderson, I., Asal, V., Rethemeyer, R.K., 2010. Lethal combinations: studying the structure of terrorist networks. In: Schmorrow, D., Nicholson, D. (Eds.), Advances in Cross-Cultural Decision Making, CRC Press, New York.
Breiger, R.L., Ackerman, G.A., Asal, V., Melamed, D., Milward, H.B., Rethemeyer, R.K., Schoon, E., 2011. Application of a profile similarity methodology for identifying terrorist groups that use or pursue CBRN weapons. Social Computing, Behavioral-Cultural Modeling and Prediction 6589, 26–33.
Dalton, A., Asal, V., 2011. Is it ideology or desperation: why do organizations deploy women in violent terrorist attacks? Studies in Conflict and Terrorism 34 (10), 802–819.
Asal, V., Young, J., 2012. Battling abroad: why some organizations are likely targets of foreign counterterrorism. Civil Wars 14 (2), 272–287.
Asal, V., Ackerman, G., Rethemeyer, R.K., 2012. Connections can be toxic: terrorist organizational factors and the pursuit of CBRN weapons. Studies in Conflict and Terrorism 35, 229–254.
Asal, V., Deloughery, K., Phillips, B., 2012. When politicians sell drugs: examining why Middle East ethnopolitical organizations are involved in the drug trade. In: Forest, J.J.F. (Ed.), Intersections of Crime and Terror. (First appeared in Terrorism and Political Violence 24, 199–212).
Melamed, D., Schoon, E., Breiger, R., Asal, V., Rethemeyer, R.K., 2012. Using organizational similarity to identify statistical interactions for improving situational awareness of CBRN activities. In: Yang, S.J., Greenberg, A.M., Endsley, M. (Eds.), Social Computing, Behavioral-Cultural Modeling, and Prediction (Lecture Notes in Computer Science 7227). Springer-Verlag, Berlin, Heidelberg, pp. 61–68.
Breiger, R.L., Melamed, D., in press. The duality of organizations and their attributes: turning regression modeling 'inside out.' In: Borgatti, S.P., Brass, D.J., Halgin, D.S., Labianca, G., Mehra, A. (Eds.), Research in the Sociology of Organizations, special issue on Contemporary Perspectives on Organizational Social Network Analysis. Emerald, London.

that they are not the best sources, and that the researcher should "talk to law enforcement people," get police records of protest events, and much more. However, these latter sources are biased as well, as discussed in Earl et al.'s (2004) review of the social movements/collective behavior literature. Moreover, writing on organized crime and terrorist groups, Kadushin (2005: 149) laments that "anyone familiar with surveillance reports knows that they [too] can be inaccurate, involve hearsay, and may reflect the biases of informants as well as those of the investigators."

With respect to selection bias, newspaper data is found to compare favorably to nonresponse in surveys (Earl et al., 2004: 77). The authors recommend triangulation of newspaper sources with other sources whenever possible. They note that a close substitute is the use of comprehensive electronic archives such as lexisNexis, which allow the search of a large number of multiple sources via keywords. In brief: researchers should avoid both the unexamined use of newspaper sources as well as blanket condemnations of their use (Earl et al., 2004).

Following from the comments above on selection bias, we reasonably expect there to be special problems with newspaper coverage of the smaller groups in our database. The same expectation is raised by Kadushin (2005) with respect to the study of terrorist groups. At a minimum, in Section A.2 we have examined the effect of different coding decisions with respect to group size. Nonetheless, the likelihood of selection bias away from coverage of the smaller groups provides yet a further caution on generalizations following from this study.

Given the established precedents of newspaper reports being used for analysis of social movements (e.g., Bearman and Everett, 1993; see also the extensive review in Earl et al., 2004) and their prior use in analyses of terrorist networks (Krebs, 2001; Morselli et al., 2007), we believe that the collection and coding of the BAAD-1 data is consistent with current best practices in this area of research.

As to the future, we both hope and expect that the standards presently maintained will rise. Gerdes and Carley (2009) point the way, by using tools of network analysis to gauge the scale of validity issues with respect to the information gap that exists between media coverage of terrorist activities in comparison to all-source analyses, finding that media sources do provide an accurate overview of the key actors involved in terrorism, but that they "lack the reliability necessary for comprehensive analysis of extremism." Gerdes et al. (2012) continue this productive line of comparison, including identifying problems in the degree of overlap of a prominent START database with other sources. We expect concerns such as these to lead to the improvement of data standards in the study of terrorist behavior and other forms of extremism.

References

Ackerman, Gary A., 2006. "Group Characteristics Database." Monteray Terrorism Research and Education Program. Monteray Institute of International Studies, Monteray, CA.

Ansell, C.K., 1997. Symbolic networks: the realignment of the French working class, 1887–1894. American Journal of Sociology 103, 359–390.

Arquilla, J., Ronfeldt, D.F., 2001. Networks and Netwars: The Future of Terror, Crime, and Militancy. RAND, Santa Monica, CA.

Asal, V., Rethemeyer, R.K., 2008. The nature of the beast: organizational structures and the lethality of terrorist attacks. Journal of Politics 70, 437–449.

Asal, V., Rethemeyer, R.K., 2009. Islamist use and pursuit of CBRN terrorism. In: Ackerman, G., Tamsett, J. (Eds.), Jihadists and Weapons of Mass Destruction. CRC Press, Boca Raton, FL, pp. 335–358.

Asal, V., Rethemeyer, R.K., Anderson, I., 2009. BAAD-1 Dataset (Big, Allied and Dangerous, BAAD). http://hdl.handle.net/1902.1/16062

Asal, V., Rethemeyer, R.K., Anderson, I., Rizzo, J., Rozea, M.M., Stein, A., 2009b. The softest of targets: a study on terrorist target selection. Journal of Applied Security Research 4 (3), 258–278.

Asal, V.H., Ackerman, G.A., Rethemeyer, R.K., 2012. Connections can be toxic: terrorist organizational factors and the pursuit of CBRN weapons. Studies in Conflict & Terrorism 35, 229–254.

Bakker, R.M., Raab, J., Milward, H.B., 2012. A preliminary theory of dark network resilience. Journal of Policy Analysis and Management 31, 33–62.

Bearman, P.S., Everett, K.D., 1993. The structure of social protest, 1961–1983. Social Networks 15 (2), 171–200.

Berry, L., Curtis, G.E., Hudson, R.A., Kollars, N.A., 2002. A Global Overview of Narcotics-Funded Terrorist and Other Extremist Groups. US Library of Congress, Washington, DC http://www.loc.gov/rr/frd/pdf-files/NarcsFundedTerrs_Extrems.pdf

Brambor, T., Clark, W.R., Golder, M., 2006. Understanding interaction models: improving empirical analyses. Political Analysis 14, 63–82.

Breiger, R.L., 1974. The duality of persons and groups. Social Forces 53, 181–190.

Breiger, R.L., 2009. On the duality of cases and variables: correspondence analysis (CA) and Qualitative Comparative Analysis (QCA). In: Byrne, D., Ragin, C.C. (Eds.), The SAGE Handbook of Case-Based Methods. Sage, London, pp. 243–259.

Breiger, R.L., Ackerman, G.A., Asal, V., et al., 2011. Application of a profile similarity methodology for identifying terrorist groups that use or pursue CBRN weapons. In: Salerno, J., Yang, S.J., Nau, D., et al. (Eds.), Social Computing, Behavioral-Cultural Modeling and Prediction. Springer, Berlin, pp. 26–33.

Burt, R.S., 1983. Studying status/role sets using mass surveys. In: Burt, R.S., Minor, M.J. (Eds.), Applied Network Analysis: A Methodological Introduction. Sage, Beverly Hills, CA, pp. 100–118.

Butts, C.T., 2009. revisiting the foundations of network analysis. Science 325, 414–416.

Byrne, D., Ragin, C.C. (Eds.), 2009. The SAGE Handbook of Case-based Methods. Sage, London.

Carley, K.M., 2003. Dynamic network analysis. In: Breiger, R.L., Carley, K.M., Pattison, P.E. (Eds.), Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. National Academy of Sciences, Washington, DC, pp. 133–145.

Casteel, S.W., 2003. Statement of the DEA Assistant Administrator for Intelligence before the Senate Committee of the Judiciary (May 20). http://www.justice.gov/dea/pubs/cngrtest/ct052003.html (accessed 1/15/2012).

Cilluffo, F., 2000. The threat posed from the convergence of organized crime, drug trafficking, and terrorism. In: Testimony before the Committee on the Judiciary, Subcommittee on Crime. US House of Representatives.

Coleman, J.S., 1988. Social capital in the creation of human capital. American Journal of Sociology 94, S95–S120.

Cornwell, B., Curry, T.J., Schwirian, K.P., 2003. Revisiting Norton Long's ecology of games: a network approach. City and Community 2, 121–142.

Correlates of War. 2004. National Material Capabilities (v3.02). http://www.correlatesofwar.org/COW2%20Data/Capabilities/nmc3-02.htm (accessed 1/2/13).

Cunningham, D., Nugent, C., Slodden, C., 2010. The durability of collective memory: reconciling the "Greensboro Massacre". Social Forces 88, 1517–1542.

de Nooy, W., 2003. Fields and networks: correspondence analysis and social network analysis in the framework of field theory. Poetics 31, 305–327.

Diaconis, P., Efron, B., 1983. Computer-intensive methods in statistics. Scientific American 248 (5), 116–130.

Diani, M., McAdam, D. (Eds.), 2003. Social Movements and Networks: Relational Approaches to Collective Action. Oxford University Press, Oxford.

DuPée, M. C., 2010. The Narcotics Emirate of Afghanistan: examining armed polities and their roles in illicit drug production and conflict in Afghanistan 1980–2010. Master's Thesis. Naval Postgraduate School, Monterey, CA.

Earl, J., Martin, A., McCarthy, J.D., Soule, S., 2004. The use of newspaper data in the study of collective action. Annual Review of Sociology 30, 65–80.

Efron, B., 1982. The Jackknife, the Bootstrap, and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics 38. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Everton, S.F., 2012. Disrupting Dark Networks. Cambridge University Press, Cambridge/New York.

Fararo, T.J., Doreian, P., 1984. Tripartite structural analysis: generalizing the Breiger-Wilson formalism. Social Networks 6, 141–175.

Faust, K., 2005. Using correspondence analysis for joint displays of affiliation networks. In: Carrington, P.J., Scott, J., Wasserman, S. (Eds.), Models and Methods in Social Network Analysis. Cambridge University Press, New York, pp. 117–147.

Felbab-Brown, V., 2010. Shooting Up: Counterinsurgency and the War on Drugs. Brookings Institution Press, Washington, DC.

Freeman, L.C., 2003. Finding social groups: a meta-analysis of the Southern Women data. In: Breiger, R.L., Carley, K.M., Pattison, P.E. (Eds.), Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. National Academies Press, Washington, DC, pp. 39–77.

George, A.L., Bennett, A., 2005. Case Studies and Theory Development in the Social Sciences. MIT Press, Cambridge, MA.

Gerdes, L.M., Carley, K.M., 2009. Estimating the accuracy of media sources on terrorism. In: Paper presented at the International Sunbelt Social Network Conference, March 12, San Diego.

Gerdes, L.M., Ringler, K., Autin, B., 2012. Examining diffusion of tacit knowledge among kidnappers in the Abu Sayyaf Group. In: Presentation at the 2012 Minerva Initiative Meeting, September 13, Arlington, VA.

Goodman, L.A., 1996. A single general method for the analysis of cross-classified data: reconciliation and synthesis of some methods of Pearson, Yule, and Fisher, and also some methods of correspondence analysis and association analysis. Journal of the American Statistical Association 91, 408–428.

Greenacre, M.J., 1984. Theory and Applications of Correspondence Analysis. Academic Press, London.

Grofman, B., Schneider, C.Q., 2009. An introduction to crisp set QCA, with a comparison to binary logistic regression. Political Research Quarterly 62, 662–672.

Harding, D.J., Fox, C., Mehta, J.D., 2002. Studying rare events through qualitative case studies. Sociological Methods & Research 31, 174–217.

Hayden, N.K., 2009. Terrifying landscapes: understanding motivations of non-state actors to acquire and/or use weapons of mass destruction. In: Ranstorp, M., Normark, M. (Eds.), Unconventional Weapons and International Terrorism: Challenges and New Approaches. Routledge, London, pp. 163–194.

Kadushin, C., 2005. Who benefits from network analysis: ethics of social network research. Social Networks 27 (2), 139–153.

Kenney, M., 2003. From Pablo to Osama: counter-terrorism lessons from the war on drugs. Survival 45, 187–206.

Krebs, V., 2001. Mapping networks of terrorist cells. Connections 24, 43–52.

LaFree, G., Morris, N.A., Dugan, L., 2010. Cross-national patterns of terrorism: Comparing trajectories for total, attributed and fatal attacks, 1970–2006. British Journal of Criminology 50 (4), 622–649.

Lauwagie, B.N., 1979. Ethnic boundaries in modern states: Romano Lavo-Lil revisited. American Journal of Sociology 85, 310–337.

Lazega, E., Jourda, M., Mounier, L., et al., 2008. Catching up with big fish in the big pond? Multi-level network analysis through linked design. Social Networks 30, 159–176.

Le Roux, B., Rouanet, H., 2004. Geometric Data Analysis: From Correspondence Analysis to Structured Data Analysis. Kluwer, Dordrecht.

Mahoney, J., Goertz, G., 2006. A tale of two cultures: contrasting quantitative and qualitative research. Political Analysis 14, 227–249.

Makarenko, T., 2004. The crime-terror continuum: tracing the interplay between transnational organised crime and terrorism. Global Crime 6, 129–145.

Marshall, M.G., Jaggers, J., Gurr, T.R., 2006. Polity IV Project: Political Regime Characteristics and Transitions, 1800–2004. Center for International Development and Conflict Management, University of Maryland, College Park, MD.

Martin, J.L., 2006. Jointness and duality in algebraic approaches to dichotomous data. Sociological Methods & Research 35, 159–192.

Martin, J.L., Wiley, J.A., 2000. Algebraic representations of beliefs and attitudes. II. Microbelief models for dichotomous belief data. Sociological Methodology 30, 123–164.

Melamed, D., Breiger, R.L., Schoon, E., 2013. The duality of clusters and statistical interactions. Sociological Methods & Research 42, 41–59.

Milward, H.B., Raab, J., 2006. Dark networks as organizational problems: elements of a theory. International Public Management Journal 9, 333–360.

Milward, H.B., Raab, J., 2009. Dark networks and the problem of Islamic jihadist terrorism. In: Goldsmith, S., Kettl, D.F. (Eds.), Unlocking the Power of Networks: Keys to High-Performance Government. Brookings Institution Press, Washington, DC, pp. 168–189.

Mische, A., Pattison, P., 2000. Composing a civic arena: publics, projects, and social settings. Poetics 27, 163–194.

Mohr, J.W., Duquenne, V., 1997. The duality of culture and practice: poverty relief in New York City, 1888–1917. Theory and Society 26, 305–356.

Morselli, C., Giguere, C., Petit, K., 2007. The efficiency/security trade-off in criminal networks. Social Networks 29, 143–153.

Murch, R.S., Tamsett, J., 2009. Early warning and prevention of jihadist WMD terrorism: law enforcement and intelligence. In: Ackerman, G., Tamsett, J. (Eds.), Jihadists and Weapons of Mass Destruction. CRC Press, Boca Raton, FL, pp. 241–258.

Nagin, D.S., 1999. Analyzing developmental trajectories: a semiparametric, group-based approach. Psychological Methods 42, 139–157.

Naim, M., 2005. Illicit: How Smugglers, Traffickers and Copycats are Hijacking the Global Economy. Doubleday, New York.

Newman, M.E.J., 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the United States of America 103 (23), 8577–8582.

Paoli, L., Reuter, P., 2008. Drug trafficking and ethnic minorities in Western Europe. European Journal of Criminology 5, 13–37.

Pattison, P.E., Breiger, R.L., 2002. Lattices and dimensional representations: matrix decompositions and ordering structures. Social Networks 24, 423–444.

Pedahzur, A., Perliger, A., 2009. Jewish Terrorism in Israel. Columbia University Press, New York.

Perliger, A., Pedahzur, A., 2011. Social network analysis in the study of terrorism and political violence. Political Science & Politics 44, 45–50.

Raab, J., Milward, H.B., 2003. Dark networks as problems. Journal of Public Administration Research and Theory 13, 413–439.

Ragin, C.C., 2000. Fuzzy-Set Social Science. University of Chicago Press, Chicago.

Ragin, C.C., 2008. Redesigning Social Inquiry: Fuzzy Sets and Beyond. University of Chicago Press, Chicago/London.

Rihoux, B., Ragin, C.C., 2009. Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Sage Publications, Thousand Oaks, CA.

Roberts, N., Everton, S.F., 2011. Strategies for combating dark networks. Journal of Social Structure 12, 1–32.

Rodriguez, J.A., 2005. The March 11th terrorist network: in its weakness lies its strength. Paper presented at the XXV International Sunbelt Conference, February 16–20, Redondo Beach, CA.

Roth, M.P., Sever, M., 2007. The Kurdish Workers Party (PKK) as criminal syndicate: funding terrorism through organized crime, a case study. Studies in Conflict & Terrorism, 901–920.

Rubinson, C., Ragin, C.C., 2007. New methods for comparative research? In: Mjøset, L., Clausen, T.H. (Eds.), Capitalisms Compared. JAI Press, Oxford, UK, pp. 373–389.

Ruggiero, V., Khan, K., 2006. British South Asian communities and drug supply networks in the UK: a qualitative study. International Journal of Drug Policy 17, 473–483.

Schaefer, D.R., 2010. A configurational approach to homophily using lattice visualization. Connections 30, 21–40.

Schweitzer, G.E., September 22, 2005. The nexus of international organized crime and terrorism: the case of dirty bombs. In: Testimony before the Subcommittee on Prevention of Nuclear and Biological Attacks, Committee on Homeland Security. US House of Representatives.

Schweizer, T., 1996. Actor and event orderings across time: Lattice representation and Boolean analysis of the political disputes in Chen Village, China. Social Networks 18, 247–266.

Siegel, D.A., 2011. Social networks in comparative perspective. Political Science & Politics 44, 51–54.

Singer, J.D., 1988. Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985. International Interactions 14 (2), 115–132.

Sparrow, M.K., 1991. The application of network analysis to criminal intelligence: an assessment of the prospects. Social Networks 13, 251–274.

Steinitz, M.S., 2002. The terrorism and drug connection in Latin America's Andean region. Center for Strategic and International Studies Policy Papers on the Americas 13. Study 5, 1–16.

Tajfel, H., 1981. Human Groups and Social Categories: Studies in Social Psychology. Cambridge University Press, Cambridge UK.

Thomas, R.G.C., 2003. What is Third World security? Annual Review of Political Science 6, 205–233.

Treverton, G.F., 2009. Approaching threat convergence from an intelligence perspective. In: Ranstorp, M., Normark, M. (Eds.), Unconventional Weapons and International Terrorism: Challenges and New Approaches. Routledge, London, pp. 141–162.

Tsvetovat, M., Carley, K.M., 2005. Structural knowledge and success of anti-terrorist activity: the downside of structural equivalence. Journal of Social Structure 6, 1–15.

United Nations General Assembly, 1999. International convention for the suppression of the financing of terrorism. Resolution 54 (December).

Wang, P., Sharpe, K., Robins, G.L., et al., 2009. Exponential random Graph (p^*) models for affiliation networks. Social Networks 31, 12–25.

Wolff, K.E., Gabler, S., 1998. comparison of visualizations in formal concept analysis and correspondence analysis. In: Blasius, J., Greenacre, M. (Eds.), Visualization of Categorical Data. Academic Press, San Diego, pp. 85–97.