Investigating the Internal Structure of Mafias Using Open-Source Intelligence

Niles Breuer

Department of Sociology

Nuffield College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Word Count: 50,999

Acknowledgements

This thesis would not have been possible without the support and guidance of all of the people who provided their assistance, time, and energy in the preparation and completion of this research.

First, I am grateful to my supervisor Federico Varese. He has been a constant supporter throughout this process. I first fell in love with the topic of criminal networks as his supervisee during my MSc and his encouragement was a major reason why I decided to pursue the DPhil. I've always appreciated our conversations about criminal networks and the nature of organized crime that pushed me to think more analytically and deeply about our topic over the years – his enthusiasm for and commitment to the field of organized crime research have been an inspiration. I am thankful to Jonathan Lusthaus for acting as my second supervisor and for always being there to offer his support. I would also like to thank Nuffield College and the Clarendon Fund for their financial support over the last three years that allowed me to undertake this project.

I would like to thank everyone who has provided their feedback, comments, and assistance on this thesis. Michael Biggs and Renaud Lambiotte acted as examiners for my Transfer of Status and Confirmation of Status, offering comments and guidance that helped shape the path I took. While not reading the thesis, I am grateful to Tomáš Diviák and Andrew Papachristos – our conversations helped me both better understand important methodological considerations and learn more about the study of criminal networks. To Martina Baradel and Elena Racheva, as well as Federico, my co-authors who worked with me on the second and third chapters in this thesis – thank you for putting up with me throughout the research and writing process. I wish to thank Jack Warburton, who tirelessly read through the endless drafts of my papers. I am also grateful to Zora Hauser, Alae Baha, Qiaoyu Luo, Fanqi Zeng, and Emilia Ziosi for their comments and feedback.

I'm grateful to my friends – Mila, Ed, Phoebe, Tom, and Christian – for their unyielding patience in listening to me constantly talk about my work and complain about the smallest things. Finally, I would like to thank my family, who encouraged me to do the DPhil and to follow my passion for criminal networks research.

Abstract

Investigating the Internal Structure of Mafias Using Open-Source Intelligence
Niles Breuer, Nuffield College
DPhil Candidate, Department of Sociology
Trinity Term, 2025

This thesis develops the use of open-source data to study the structure of mafias. The thesis presents two empirical papers that investigate the structure of inter-group coordination within mafias and two empirical papers that focus on how open-source business register data can be used by both academic researchers and law enforcement to investigate the structure of mafias as they infiltrate legal markets.

The first paper, *Investigating the Dynamics of Yakuza Violence Using Multilevel Network Analysis*, examines a multilevel network of violent yakuza-on-yakuza attacks constructed using media sources. New multilevel temporal reciprocity measures and a multilevel exponential random graph model find that yakuza dispute resolution mechanisms are effective at reducing the complexity of violence networks and that yakuza groups identify themselves with the national syndicate to which they are affiliated. The second paper, *Factions and Brokers in the Russian Mafia: Investigating the Structure of the Thieves-in-Law*, investigates a co-signing network of documents published and signed by *vory*, the bosses of the Russian Mafia, finding that the formally non-hierarchical *vory* form factions around vertical patronage relationships rather than ethnicity or age, with key brokers bridging otherwise-disconnected clusters. The third paper, *Testing the Reliability of OSINT Network Data for Investigating Organized Crime Infiltration of Legal-Market Businesses*, compares data from the Italian business register to a closed-source indictment from a case concerning an `Ndrangheta group. Results suggest that while open-source business register data provide unique benefits, researchers must be careful to apply the data to analyses for which they are well-suited. The fourth paper, *Proxy Targeting and Multiplex Resilience for Organized Crime Network Disruption*, uses the business register and indictment datasets to investigate whether open-source data can be useful for targeting surveillance and disruption of mafia networks in legal markets. The analyses suggest that open-source data offer a useful proxy for expensive and hard-to-collect closed-source data when targeting criminal network structures in legal markets; however, tie multiplexity reduces the impact of targeted disruptions.

Contents

List of Figures

List of Tables

1

Introduction

Despite the illegality of their activities, organized crime groups are organizations like any other: "[Illegal enterprises] can be studied as businesses, and the same questions can be asked that would be asked of analogous legal businesses" (Haller, 1990, p. 208; see also Gambetta, 1993). Organized crime groups – like corporations, social associations, and governments – must recruit members, govern internal processes, and coordinate to carry out tasks. These organizations, both legal and illegal, face challenges to these tasks: how to recruit effective and loyal members under conditions of imperfect information, how to maintain internal order, how to ensure cooperation between actors and organizational sub-units with distinct priorities and goals (Barnard, 1938; Catino, 2019). Theories and insights into how people communicate and collaborate and how organizations are structured should extend beyond legal organizations to also cover those involved in criminal activities (Catino, 2019). At the same time, organized crime groups operate under unique constraints – illegality means that these groups cannot appeal to the legal system to resolve disputes and must actively hide their operations from that legal system (Catino, 2019; Reuter, 1983; D. C. Smith, 1971). This means that how organized crime groups solve particular challenges can contribute to our understanding of organizations more broadly, like the role of multiplexity, the relationship between formal organizational structures and the informal connections that members form amongst themselves, and the relationships of subgroups within a larger multilevel divisional organization.

Within the broad category of organized crime, mafias are particularly well-suited for sociological and organizational theory research. Mafias are criminal groups defined by their provision of private protection services and aim of governing markets and communities (Breuer & Varese, 2023; Catino, 2019; Gambetta, 1993; Paoli, 2004; Varese, 2001, 2013a).[1] Prototypical examples of mafias include the Sicilian Cosa Nostra, the Calabrian `Ndrangheta, the Italian-American Mafia, the Japanese yakuza, the Hong Kong triads, and the post-Soviet *vory*. These mafias are longstanding – having endured for decades or centuries – and have survived multiple leadership successions (Paoli, 2020; Reuter & Paoli, 2020). They are also organizationally complex, typically exhibiting multiple levels and divisions into subgroups, with both formal hierarchies and informal systems based on social relationships and kinship (Catino, 2019; Kaplan & Dubro, 2012; Paoli, 2020). These factors mean that a range of organizational aspects can be studied by investigating mafias. In this thesis, I focus on the informal structure of mafias that arises from the communication, collaboration, and relationships between mafiosi within their organizations.

The academic literature on this aspect of mafias has grown in the last twenty years. In particular, social network analysis techniques have been used to map out the relationships

[1] While mafias are distinguished from other organized crime groups by their focus on protection and governance, mafia groups and individual mafiosi can become involved in other illegal activities like illegal gambling, loansharking, drug trafficking, and infiltration of legitimate businesses (Calderoni, 2012; Paoli, 2020; Reuter, 1983; Reuter & Paoli, 2020).

between mafiosi and disentangle the mechanisms that drive mafiosi to form connections with each other. Prior research has primarily focused on three Italian mafias based in Italy and the United States – the Sicilian Cosa Nostra, the Calabrian `Ndrangheta, and the Italian-American Mafia – and on a small set of criminal activities, mainly the provision of governance and the trafficking of cocaine in the case of the `Ndrangheta (Calderoni, 2012, 2014, 2016; DellaPosta, 2017, 2023; Krajewski et al., 2022; Mastrobuoni, 2015; Mastrobuoni & Patacchini, 2012; Musotto, 2022). While the structure of other mafias has been investigated to a lesser extent, the evidence has largely been anecdotal and inconclusive (Cowley et al., 2015; Galeotti, 2018; Hill, 2003; Kaplan & Dubro, 2012; Slade, 2014).

This focus on a small set of mafias and illegal activities is driven in part by the availability of data: while other fields within sociology and other social sciences regularly make use of a wide range of often publicly-available data sources, the study of the structure of mafias – as well as other organized crime groups – has mostly relied on closed-source law enforcement and judicial files given the nature of the social phenomena under investigation (Bright et al., 2021). Thus, the groups and activities that are studied are those for which such data are most easily accessible, leaving other mafias understudied. This leaves a gap in our understanding of organizational dynamics within mafias – we know little about non-Italian mafias and about other activities they undertake. Only considering the mafias and criminal activities for which we have official judicial or police data limits our ability to draw broader conclusions.

This thesis aims to address this limitation by investigating understudied mafias and mafia activities using social network analysis. To do so, I use a range of open-source and publicly-available datasets to enable the investigation of how different mafias and different criminal activities are structured. While high-quality closed-source data is rare, other less commonly used sources provide an opportunity to expand the study of mafias. In the following papers, I focus on high-level coordination within the yakuza and the *vory* and on the involvement of the `Ndrangheta in legal markets. These papers focus on investigating how these mafias are structured, what the determinants of those observed structures are, and the impact that they have on the operations and resilience of the organizations. Further, I contribute to the study of mafias and other organized crime groups by developing new data sources, methodological considerations, and analytical techniques in the course of these papers that can be applied to other criminal groups.

This articles-format thesis consists of four standalone but thematically-connected papers, along with a general introduction and conclusion that frame these individual chapters. Two of the papers are co-authored, while two are sole-authored: the first paper has been co-authored with Dr. Martina Baradel, while the second paper has been co-authored with Prof. Federico Varese and Dr. Elena Racheva. In the case of the co-authored papers, I am the first author and led the conceptual development, data cleaning and preparation, data analysis and interpretation, and manuscript drafting. The papers in this thesis have been adapted from four research articles for which I drafted the manuscripts:

1. "Investigating the Dynamics of Yakuza Violence Using Multilevel Network Analysis", published in *Social Networks*, 2025.
2. "Factions and Brokers in the Russian Mafia: Investigating the Structure of the Thieves-in-Law", published in *British Journal of Criminology*, 2025.
3. "Testing the Reliability of OSINT Network Data for Investigating Organized Crime Infiltration of Legal-Market Businesses", published in *Global Crime*, 2025.
4. "Proxy Targeting and Multiplex Resilience for Organized Crime Network Disruption in Corporate Interlock and Communications Networks", published in *Social Network Analysis and Mining*, 2025.

1. Introduction

Full bibliographic information is provided on the title page of each chapter.

The first and second papers focus on coordination within mafias above the level of the main operational units; rather than focusing on individual sub-units or 'families', I focus here on how these sub-units and their bosses interact and on how the patterns of these informal interactions converge with or diverge from the formal structure of the organization. The first paper in this thesis, *Investigating the Dynamics of Yakuza Violence Using Multilevel Network Analysis*, examines the multilevel structure of Japanese yakuza syndicates based on violent conflict between their subgroups. Research on mafia conflict has mainly focused on how internal systems are used to curb violence (Catino, 2014, 2019; Thompson, 2024). However, little is known about the patterns of violence among mafia groups and organized crime groups. Further, prior ethnographic research on the yakuza has ascribed varying levels of control over affiliated groups to yakuza syndicates – some accounts suggest that these syndicates wield authority over their affiliated groups, while other accounts report that yakuza groups are largely autonomous and syndicates are only involved in decision-making relating to issues that could potentially impact other syndicate members beyond a group's own territory (Gragert, 1997; Hill, 2003; Kaplan & Dubro, 2012). To investigate these patterns of mafia conflict and what they suggest about the internal organizational structure of the yakuza, we collected a novel dataset of violent incidents between local yakuza groups across Japan from publicly-available media sources for the six-year period between 2014 and 2019; these data were used to construct a multilevel network that represents the relationships between yakuza syndicates, their affiliated local groups, and the cities and prefectures in which those local groups are headquartered. We defined new multilevel temporal reciprocity measures and used a multilevel exponential random graph model to understand the drivers of yakuza-on-yakuza attacks.

1. Introduction

We find support for the claim that yakuza syndicates function as cohesive organizations with a shared identity and the ability to constrain the actions of their constituent groups – groups affiliated to the same syndicate tend not to attack each other and tend to retaliate on each other's behalf. Crucially, we find that this retaliation typically takes the form of attacking another group affiliated with an opponent's syndicate rather than the opponent group itself, further supporting the claim that yakuza groups identify themselves and other groups with their associated syndicates. Alliances between syndicates, on the other hand, do not create a sense of shared identity between the groups of member syndicates: while alliances are successful at reducing conflict internally, groups do not retaliate on behalf of their allies. These results clarify the complex organizational structure of the yakuza – while local groups may be the main operational units of the yakuza, these groups identify with and are constrained by the syndicates to which they are affiliated. Further, the structure of the yakuza provides conflict resolution infrastructure that enables groups and syndicates to negotiate settlements to disputes before they can cascade into complex patterns of violence. This paper stresses the importance of considering the multilevel divisional structure of mafias – looking only at the local yakuza groups and not considering the syndicates to which they belong obscures the multilevel considerations driving attacks. Further, these results suggest that the formal syndicate-based hierarchy and the informal structure of the yakuza – based on the decisions of groups to attack one another – converge, with syndicate membership creating a bond between groups that is not seen amongst gang consortia in North America (Bichler et al., 2019; Decker & Curry, 2002; Randle & Bichler, 2017).

While prior research has investigated the structure of mafia groups and families as they execute their criminal operations, little is known about how mafia bosses interact with one another and the structure of their higher-level coordination bodies. The second paper, *Factions and Brokers in the Russian Mafia: Investigating the Structure of the Thieves-in-Law*, contributes to the literature on mafia structure by presenting the first direct analysis of the structure of a mafia commission. We focus on the *vory-v-zakone* or 'thieves-in-law', the bosses within the constellation of post-Soviet organized crime groups colloquially known as the Russian Mafia. The fraternity of these bosses is formally egalitarian and non-hierarchical – officially, all *vory* are equal and have the same power and authority. However, prior ethnographic research suggests that the *vory* have informal power hierarchies (Cowley et al., 2015; Handelman, 1995). Further, some accounts suggest that the *vory* association has fragmented into separate subgroups based on ethnic and generational lines (Mironova, 2023; Slade, 2014). To investigate the structure of this higher-level coordination body, we collected a dataset of *progony* – documents published and signed by the *vory* – to construct a co-signing network representing which mafiosi signed documents together. We analyze the structure of this network using both descriptive and simulation-based network analysis techniques to investigate the underlying social system represented by decisions of *vory* to co-sign together.

We find that the *vory* exist within a polycentric association that exhibits signs of factionalization, with *vory* clustering around different well-connected and influential actors. Well-connected nodes tend not to be connected to each other; when they do connect, however, they tend to form strong ties. Factions are connected by a small number of key brokers that bridge between otherwise-disconnected components of the network. Rather than factions arising along ethnic and generational cleavages leading to fragmentation as suggested by prior research, however, we find that factions are more associated with cross-cutting vertical patronage relationships that connect *vory* of different ethnicities and ages. These findings provide a novel view inside of the *vory* fraternity and have implications for other mafia commissions. As mafias elsewhere in the world like Italy, the United States, and Japan have come under increasing pressure, they have struggled to maintain their higher-level coordination bodies in the face of mounting coordination problems. Further, while the case of the yakuza shows signs of convergence between formal and informal structures, the *vory* exhibit the opposite – while formally egalitarian and non-hierarchical, informal relationships amongst the mafiosi reveal fragmentation and power structures. This paper contributes to the literature of mafia structure by directly investigating the drivers of relationships between mafia bosses.

Beyond mafias that have been understudied, the existing literature has also failed to investigate the structure of mafias as they carry out different criminal activities. The third and fourth papers in this thesis focus on how mafias are structured as they seek to infiltrate legitimate businesses. While this topic has become more popular in recent years, the academic literature has largely focused on what sorts of businesses are infiltrated (Mirenda et al., 2019; Sciarrone & Storti, 2014), what impact this has on local economies (Barone & Narciso, 2015; Checchi & Polo, 2020), and how infiltrated businesses can be identified (Aziani et al., 2022; Jofre, 2022; Jofre et al., 2024). Little is known about the structure of mafias in legal markets. This gap in our understanding is due, in part, to a lack of reliable and rich data on the ties between mafiosi as they look to found or join legitimate companies. The third paper in this thesis, *Testing the Reliability of OSINT Network Data for Investigating Organized Crime Infiltration of Legal-Market Businesses*, aims to address this limitation by evaluating whether a promising data source – open-source business registers – can be used to study this phenomenon. To do so, I test whether publicly-available business register data (1) can be helpful when investigating the structure of mafias and other organized crime groups in legal markets and (2) can serve as a reliable proxy for harder-to-access closed-source data. The first research aim concerns the strengths and weaknesses of open-source and closed-source data, while the second seeks to understand how much is lost due to the fact that these data cannot contain information on hidden criminal activities.

Using data from the Italian business register – Il Registro Imprese – and a pre-trial detention order relating to a case involving an `Ndrangheta group and its associates infiltrating legitimate businesses in Italy and Germany, I construct four networks representing different relationships between the involved actors. There are two corporate interlock networks, with one representing which actors are officially involved in the same companies based on the business register and the other representing which actors are officially or covertly affiliated with the same businesses based on the pre-trial detention order; further, there are two indictment-based communication networks representing general criminal communications and business-related communications respectively. Comparing the strengths and weaknesses of the publicly-available and closed-source data, I find that business register data can be useful for academic researchers and law enforcement analysts to answer certain research questions, but certain analyses must be done with caution due to concerns about reliability. Business register-based networks are easily accessible, inexpensive, and rich in data on actors and companies; further, these data enable research on temporal and post-disruption structural dynamics that cannot be studied using only data from commonly-used closed sources. Business registers can also act as supplements to closed sources: based on the business register data, I identify that the pre-trial detention order contained incomplete or erroneous ownership information for 13% of the relevant companies. Finally, I find that the business register-based network does not reliably represent the node-level or global structural characteristics of either the indictment-based corporate interlock network or the communication networks. The observed dissimilarity in node-level and global metrics, however, obscures the similarity of the business register-based and indictment-based networks at the local neighborhood-level. While researchers should be cautious when ranking actors by centrality scores or calculating graph-level measures based solely on publicly-available data – since these analytical levels are more susceptible to distortion from the exclusion of covert ties – these results suggest that business register data are suitable for investigating the underlying tie formation mechanisms in networks using exponential random graph models, stochastic actor-oriented models, and other network models that investigate how actors form ties based on local network configurations. While some caution is warranted, publicly-available business register data present an opportunity to investigate the structure of mafias and other organized crime groups as they infiltrate legal markets. This paper demonstrates the utility and limitations of a novel data source and suggests that other potential data sources should be evaluated in a similar manner.

Building on the third paper, the fourth paper – *Proxy Targeting and Multiplex Resilience for Organized Crime Network Disruption* – proposes and implements a novel approach to network disruption analyses. Identifying strategies for disrupting mafias and other organized crime groups has been one of the major priorities for academic research on criminal networks. One prominent approach to this topic has been to simulate the effect of removing nodes or edges from network objects that represent mafia structures to understand how this disruption impacts the network's ability to function (Agreste et al., 2016; Bright et al., 2014; Cavallaro et al., 2020; Wood, 2017). This approach, however, was developed for monoplex networks, where ties between actors represent only one type of social relationships. Prior research, however, suggests that mafias and organized crime groups are multiplex, with members connected by a range of different relationship types in different dimensions like collaboration, communication, kinship, and friendship (Diviák et al., 2019; Malm et al., 2010; C. M. Smith & Papachristos, 2016). Further, prior disruption analysis approaches have typically assumed complete knowledge of the underlying network; this information, however, is often not available to law enforcement until late into an investigation, limiting the practical utility of these analyses (Manzi & Calderoni, 2024b; Musciotto & Miccichè, 2022). This paper proposes a novel approach to network disruption analyses that better accounts for both multiplexity and the sort of information available to police during an investigation. Rather than selecting nodes for removal from a network based on their position in that same network, the proxy targeting approach laid out in this thesis separates the targeting and removal steps: nodes' positions in one dimension of a multiplex network are used to determine which node to remove from a different dimension in the network. The proxy targeting approach is applied to the network datasets presented in the third paper, conceptualized here as different dimensions in a multiplex network.

The results in the fourth paper suggest that business register network data can be an effective and easy-to-access means for targeting surveillance and disruption of mafias as they look to infiltrate the legal economy; however, the communication dimension between relevant actors is less impacted. Further, findings suggest that multiplexity can increase the resilience of a network. Rather than key actors being important for connectivity across multiple dimensions, in this case removing central actors from one dimension has a minimal impact on the cohesion of other dimensions. However, this weakens as more nodes are removed from the network: while multiplexity may reinforce against small disruptions, it may not prevent destabilization in the face of greater shocks. The analyses in this paper suggest that multiplexity is a key consideration when designing strategies to disrupt networks or designing organizations to withstand destabilization. This study contributes to the literature by extending disruption analyses to the understudied topic of mafia activities in legal markets, developing and testing a new analytical approach, and examining the importance of multiplexity to the resilience of criminal networks.

Finally, the Conclusion summarizes the results of the four preceding papers and discusses their criminological and sociological implications. I conclude by discussing avenues for future research.

2

Investigating the Dynamics of Yakuza Violence Using Multilevel Network Analysis

Abstract[2]

This paper investigates the structure of the yakuza – the Japanese mafia – and the patterns of violence between local yakuza groups using a novel dataset of yakuza-on-yakuza conflict throughout Japan between 2014-2019. We define new multilevel temporal reciprocity measures and apply a multilevel exponential random graph model to investigate the structure of yakuza violence. We find low levels of retaliation and complex 'cascading' conflict structures and that yakuza syndicates act as cohesive organizations that can constrain the actions of their member groups. This research contributes to the understanding of the yakuza's structure and how violent conflict occurs within organized crime groups.

[2] This work was done in collaboration with Dr. Martina Baradel. A version of this paper has appeared in: Breuer, N., & Baradel, M. (2025). Investigating the dynamics of yakuza violence using multilevel network analysis. *Social Networks*, *82*, 14-26. https://doi.org/10.1016/j.socnet.2025.03.001

2.1 Introduction

Violent conflict is a key way for organized crime groups to compete with one another in the course of their criminal operations. Prior research on gang violence and mafia conflict has explored the reasons behind disputes and patterns of violence between clashing groups, with mixed results. Research on mafia conflict has mainly focused on how internal systems are used to curb violence (Catino, 2019; Thompson, 2024). However, little is known about the *patterns* of violence among mafia groups and organized crime groups. While these patterns have been investigated for street gangs, the organizational structure of these groups has been found to lead to specific forms of conflict that may not generalize to organized crime groups like mafias (Decker et al., 2008; Decker & Curry, 2002; Lewis & Papachristos, 2020).

This article looks at the case of the Japanese yakuza to analyze patterns of violence in organized crime groups' conflicts. Japan offers a particularly clear case for understanding these patterns of conflict because, due to their recognized status[3], the yakuza presents one of the highest levels of hierarchical sophistication and internal systems of dispute resolution. We use newly-defined multilevel temporal reciprocity measures and a multilevel exponential random graph model (ERGM) to analyze a novel dataset of yakuza attacks throughout Japan from 2014-2019. In-line with prior theories about conflict resolution in organized crime groups, we find that the national-level organizational layer of yakuza syndicates, high levels of inter- and intra-syndicate coordination, and the formality of alliances create opportunities for yakuza groups to resolve violence using methods that gangs and even other mafias cannot. While syndicates function as cohesive organizations, alliances between syndicates are effective at

[3] Yakuza membership is not criminalized in Japan.

reducing conflict between allies but fail to function as unified institutions. Further, we find that while syndicates are cohesive, local factors still drive conflict between groups. More counterintuitively, we find that retaliatory violence is not straightforward in the majority of cases: yakuza groups do not strike back directly when attacked. Instead, yakuza groups engage in indirect syndicate retaliation, where a group from the same syndicate retaliates on the original victim's behalf against another group from the syndicate of the original attacker.

By looking at how yakuza groups respond when attacked, we contribute not only to the understanding of the yakuza, but to the field of mafia conflict and organized crime group violence. Our yakuza violence incident database, novel analytical framework, and new network analysis measures advance previous research on gang violence and mafia conflict by expanding from analysis of purely local factors to a combination of local and nation-wide dynamics and accounting for the multilevel nature of violent conflict. This novel dataset and approach allow for direct investigation of the dynamics of conflict between mafia groups, an aim that has previously been difficult to realize. Our findings also have implications for how yakuza violence in Japan and mafia violence more generally is likely to change as law enforcement agencies continue to pressure these criminal organizations. Finally, this research demonstrates the utility of using publicly-available open-source data when investigating organized crime groups, rather than relying solely on data from difficult-to-access police or judicial files.

The article is structured as follows. Section 2.2 reviews the relevant literature on gang violence, mafia conflict, and the structure of the yakuza and presents hypotheses on the structure of yakuza violence. Section 2.3 introduces the data and methods. Section 2.4 presents the main results. Section 2.5 discusses the results, while Section 2.6 concludes by discussing the implications of our analysis. Japanese terms without a direct translation – especially yakuza-related terms – are disambiguated in the footnotes.

2.2 Theoretical Background and Hypotheses

2.2.1 Literature Review

The literature on patterns of violence between criminal groups has largely focused on violence among street gangs in North America, particularly the United States. Within this context, gangs are portrayed as "loosely organized confederations" that lack the organizational cohesion and institutional structures necessary to control the actions of their members (Decker & Curry, 2002, p. 351; see also Decker et al., 2008). The decision of a gang member to attack members of another group, or their own group, is one aspect that North American gangs are thought to be unable to control. This literature has primarily centered on status considerations as a central driver of gang violence patterns, with 'saving face' or bolstering reputation as key motivations for committing violent attacks (Gravel et al., 2018; Hughes & Short, 2005; Lewis & Papachristos, 2020; Nakamura et al., 2020; Papachristos, 2009). Competition over territory and resources also provides opportunity and incentive for violence (Bichler & Norris, 2024; Randle & Bichler, 2017). While findings have generally been mixed, the existing literature has typically found evidence of reciprocity, where gangs tend to launch a retaliatory attack after being victimized (Hughes & Short, 2005; Lewis & Papachristos, 2020; Papachristos, 2009; Papachristos et al., 2013); retaliatory violence operates as a "rule of the game" where victims must strike back to avoid losing status (Lewis & Papachristos, 2020, p. 1830). Gang violence also often leads to complex patterns involving three or more gangs, such as chains – where the first victim does not strike back against their original attacker but instead attacks an unrelated gang – and stars where multiple gangs attack the same victim or one gang attacks multiple victims (Randle & Bichler, 2017). Some studies find evidence of more complex patterns like transitive triads or generalized exchange (Bichler et al., 2020), although findings are mixed (Papachristos et al., 2013). These network structures represent cascading violence, where an initial attack overflows beyond the initial attacker and victim. Two gangs having overlapping or adjacent territory increases the likelihood of conflict, as spatial proximity increases opportunity for attacks and can give rise to territorial conflicts or expansion by one gang into the turf of another (Papachristos, 2009; Papachristos et al., 2013; Tita & Radil, 2011).

These trends are susceptible to external influences, such as the action of law enforcement, as seen in the reaction of gangs to civil gang injunctions in Los Angeles. Civil gang injunctions are regulations that restrict the behavior of gangs or affiliated members in designated areas, with the aim of suppressing social interactions likely to spark or increase violence. Gangs that have been enjoined under a civil gang injunction grew more aggressive towards other gangs and became enmeshed in denser networks of violence with more complex structural patterns (Bichler et al., 2019; Bichler et al., 2020, 2021). The increased frequency of attacks and complexity of conflict relationships result from both status considerations and material needs: the imposition of a civil gang injunction could compel the recipient both to reassert its status through violence after being made to look weak and to seek new areas in which to conduct criminal business, leading to conflict over resources and territory. Some gangs are also impacted by multilevel factors by being part of gang consortia somewhat analogous to yakuza syndicates. The most prominent of these consortia are the Crips and the Bloods, two rival gang factions that grew out of Los Angeles in the 1960s and 1970s and subsequently spawned and recruited affiliated gangs throughout the United States. While gangs within a consortium may share a linked identity and some cultural aspects, these gang superstructures do not act as cohesive organizations and lack the ability to control the behavior of their members: prior research has confirmed high levels of intra-consortium violence, with gangs attacking others affiliated to the same faction (Bichler et al., 2019; Decker & Curry, 2002; Descormiers & Morselli, 2011; Randle & Bichler, 2017).

While the disorganized structure of street gangs leads them to engage in status-focused violence, many mafias and some other organized crime groups have developed higher-level coordination bodies to mediate conflict and make decisions about the use of violence. Extending Schelling (1980)'s argument on the strategy of conflict to organized crime groups, pure conflict, in which two opponents have completely opposite interests and therefore stop fighting only when one has exterminated the other, is quite rare. In most cases, conflict is solved by bargaining, accommodating, and avoiding damaging behavior to reach an outcome that is mutually advantageous. Operating in a world where conflict cannot be mediated through the courts of law, criminal organizations must resort to alternative methods of dispute resolution. In some cases, mediation can be provided by institutions external to the organization: for instance, Somali pirates solve their disputes through clan elders, who refer to the Somali customary law (*xeer*) and Sharia law to regulate social interactions (Shortland & Varese, 2016).

Some mafias have instead developed internal systems for coordinating action and resolving internal disputes; these higher-level coordination bodies are made up of the basic operational units of the mafias like 'families' (Catino, 2019). These bodies create mechanisms for communication, coordination, and dispute settlement (Thompson, 2024). One of their key functions is to regulate the use of violence: while being a key resource for mafias, violence is expensive as it can catalyze internal conflict or repression from external sources (Catino, 2014; Gambetta, 1993). The Sicilian Cosa Nostra and the `Ndrangheta both possess similar vertical higher-level coordination bodies that centralize power over certain decisions above the level of the otherwise-autonomous families. These systems are made up of multiple layers of aggregation from families into provincial commissions that culminate in the regional commission, which is vested with the authority to set rules for all constituent families and mediate disputes (Catino, 2019). In recent years, after the arrests of a number of prominent bosses, this structure has fallen apart and was replaced by a more horizontal system of interaction among groups from different provinces (DIA, 2020). Catino (2014) finds that mafias with higher-level coordination between families – the Sicilian Cosa Nostra and the `Ndrangheta – experience lower rates of conflict and homicides and more high-profile assassinations than the more distributed Camorra, which is composed of numerous clans organized horizontally without a unifying coordination system similar to gang consortia. The Cosa Nostra and the `Ndrangheta's vertical superstructures acted to constrain the actions of member families and reduce inter-family conflict while allowing these mafias to identify and target common external enemies to respond against state repression (Catino, 2014). The Camorra, on the other hand, experiences more frequent bursts of inter-clan violence and internal leadership struggles (Catino, 2019; Gambetta, 1993; Massari & Martone, 2018).

When mafia conflict does occur, violence can cascade between groups and territories. As mafias expand to new territories, there might be a spillover effect in which conflict in the territory of origin is moved to new non-traditional territories where mafias operate. In Italy, conflict in mafia strongholds reflects in the pattern of violence in territories where mafias have moved. The intensity of their presence influences the levels of violence in which they engage:

the more entrenched in the new territory, and the more capacities they developed to inflict violence, the more likely it is that conflict will be transferred to the new area (Moro & Sberna, 2018). Mafia conflict has also shown a tendency towards complex structures like those seen in some studies of North American gang violence networks, with an initial attack leading to cycles of retaliatory violence within a mafia (Neumann et al., 2017).

2.2.2 Yakuza Background and Hypotheses

Japanese crime syndicates, known as yakuza or *bōryokudan*, include 25 syndicates with a total of 23,000 members divided into local chapters or groups across Japan (NPA, 2022). They are considered a mafia, a form of governance that attempts to supply private protection (Varese, 2017). Operating shoulder to shoulder, competition for resources and territory has often led to conflict. However, when compared to other parts of the world in which multiple criminal organizations operate, the yakuza have always been successful in maintaining low levels of violence and, notably, have hardly ever caused casualties among civilians and law enforcement officers (Catino, 2019).

Similar to other mafia groups, yakuza groups are structured hierarchically. The internal structure of a yakuza group takes the form of a fictive family and is based on vertical relationships between the boss and his affiliates, and horizontal and diagonal relationships between members.[4] The local groups are the primary operational units of the yakuza (Hill, 2003; NPA, 2007). Yakuza groups are the basic building blocks of yakuza syndicates, which typically take the shape of nested, multi-layered pyramids: individual local groups control territory within a city and are affiliated to a higher-layer group made up of multiple local groups.

[4] The groups are referred to as *ikka* or *gumi*. The boss is referred to as *oyabun* (father figure) and the affiliate as *kobun* (son figure).

At the top of the standard pyramidal structure is the boss of the syndicate who is surrounded by a number of advisors and underbosses.[5] The Rokudaime Yamaguchi-gumi, the largest yakuza syndicate, contains five layers within its pyramidal structure (Jitsuwa Jidai Henshūbuhen, 2002). These national syndicate structures – each made up of a number of local groups united under the syndicate's banner – maintain relationships with each other, sometimes formalized as official alliances between syndicates.

We represent this multi-layered system with a multilevel network, illustrated by the conceptual visualization in Figure 2.1. This system contains four levels of nodes: local yakuza groups, yakuza syndicates, cities, and Japanese prefectures. The main units of analysis are local yakuza groups. Directed edges between groups represent violent attacks. Each of these groups is then tied to the syndicate to which it is affiliated; further, each group is headquartered in one city. Edges between syndicates represent formal alliance agreements. Edges between cities represent cases where two cities are geographically contiguous with one another; each city is then tied to the Japanese prefecture in which it is located. These prefectures are similarly tied when they are adjacent.

Prior accounts of the yakuza have ascribed varying levels of control over affiliated groups to yakuza syndicates. Gragert (1997) described yakuza groups as "theoretically […] autonomous" but notes that once they have become affiliated with a syndicate, they are subject to syndicate policy as set by the organization's boss and executives and are required to send payments[6] each month to the syndicate (Gragert, 1997, p. 164). Hill (2003) reports that syndicates only exert control over their affiliated groups on issues that could potentially impact other syndicate members beyond a group's own territory, such as leadership succession or

[5] Executives: *kanbu;* the boss of the syndicate: *kumichō;* advisors*: kōmon; wakagashira.*
[6] *Jōnōkin* or *kaihi.*

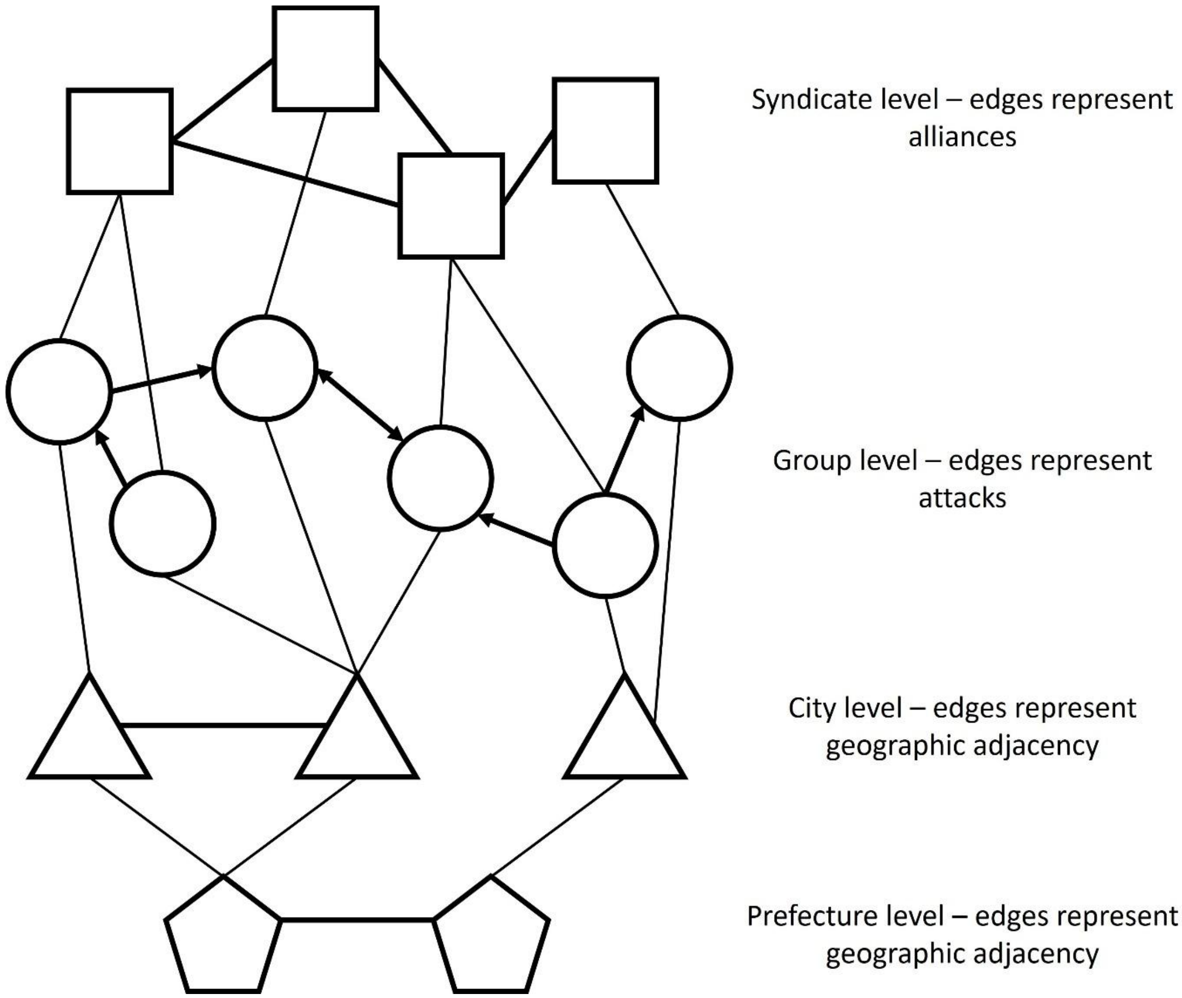


Figure 2.1. Multilevel network structure, including alliances between syndicates, local groups' affiliations to syndicates, attacks between local groups, local group headquarters in cities, geographic adjacency between cities, cities' affiliations to prefectures, and geographic adjacency between prefectures.

plans to enter new territory; outside of decisions with far-reaching implications, however, he argues that local groups function largely autonomously and make day-to-day decisions relating to business and organizational operations. Comparing the yakuza to other mafias in the United States and China that feature similar structures of basic operating units and higher-level coordination bodies, Kaplan and Dubro (2012) describe yakuza syndicates as more organizationally complex institutions with "enormous control vested at the top" (Kaplan & Dubro, 2012, p. 117).

Based on these prior accounts, we expect that yakuza syndicate structures – unlike gang consortia – will act as cohesive organizations that constrain the actions of their constituent members. As in the case of mafia groups in Italy, we expect high levels of coordination within syndicates to allow for the mediation of disputes between member groups and the early resolution of violent conflicts before they lead to a complex cycle of violence (Catino, 2014). Further, we expect yakuza groups to interpret an attack on a member of the same syndicate as an attack on their overall organization and to retaliate against the original attacker or another group within the attacker's syndicate – if yakuza syndicates act as unified bodies, then reprisal against the original attacker's syndicate can serve the same purpose as direct retaliation. Therefore, we expect:

> H1a. Yakuza groups are less likely to attack other groups affiliated with the same syndicate.
> H1b. Yakuza groups are less likely to retaliate against other groups affiliated with the same syndicate.
> H1c. Yakuza groups are more likely to retaliate on behalf of other groups affiliated with the same syndicate.

Syndicates also entertain relationships with each other. Alliances between syndicates help prevent conflict over territory and increase collaboration; agreements may work on a temporary basis or develop into longstanding positive relationships (Gragert, 1997). The three main alliances are the Kantō Shinbokukai in the Kantō area, the Goshakai in the Chūgoku and Shikoku areas, and the Yonshakai in Kyūshū. Alliances typically agree in advance on punishments for transgressors within the alliance structure, such as temporary or permanent expulsion for a person responsible for an attack on a member group's boss or for harming civilians or law enforcement officers. Syndicates also agree to cooperate against aggressors external to the alliance. Alliances are maintained through regular meetings of the executives from member syndicates, who discuss business matters and conflict reduction; groups within each member syndicate are then notified of the meeting's decisions. Alliances can be formed between individual syndicates through the exchange of sake cups and the participation in other syndicates' ceremonies. For instance, while the Yamaguchi-gumi is not part of the Goshakai alliance, the involvement of the Yamaguchi-gumi's ranked members as guarantors in Goshakai's ceremonies indicates a positive connection between the syndicate and the members of the alliance (JJ11/2014). These alliances are reported to be effective in the resolution of conflict. As the ex-legal counselor of a main syndicate reported: "The yakuza cannot appeal to the law because they abide by a different set of norms, so these alliances work as an alternative dispute resolution system based on yakuza rules. The relationships within the alliance are sanctioned by a formal ritual, therefore they are very strict and they are respected" (Interview S3).

Syndicates and alliances provide channels for resolving disputes. For instance, syndicates may impose a 'no return' rule that prohibits groups from striking back when attacked in order to avoid escalating the conflict (JJ10/2016). On some occasions, the groups involved proceed directly to the settlement before there is any time to strike back

(JJ07/2007). Especially since the implementation of stricter anti-yakuza laws, the resolution system has been swift: the Azabu conflict, involving the Yamaguchi-gumi and Sumiyoshi-kai syndicates that are allied with each other, started when a member of the Sumiyoshi-kai was shot dead in his car on February 5th 2007. Within 20 hours, four attacks followed rapidly but then they suddenly stopped; both syndicates had circulated a standby order that prohibited revenge attacks, while top management met and agreed not to escalate. Three days later, the two groups settled the matter officially, both agreeing that an escalation of the conflict would have been detrimental for the groups (JJ04/2007; JJ07/2007).

Mediation of disputes typically occurs through a reconciliation ceremony, the result of which is made public knowledge. In this ceremony, after a mediator has secured settlement offers from the parties in the dispute – often involving financial compensation – an arbitrator and the parties meet in neutral territory and make a public and formal agreement to end the dispute, with the explicit provision that further attacks related to this dispute would cause the members of the ceremony's audience to turn against the aggressor (Jitsuwa Jidai Henshūbuhen, 2007a). The formal and public ceremony makes the agreement easily enforceable, as deviation is easy to monitor and the large number of witnesses increases the cost of defection (Hill, 2003), and until such ceremony takes place, attacks are still possible (see Jitsuwa Jidai Henshūbuhen, 2007b, p. 217). The settlement necessary for this ceremony is easier for yakuza groups to reach when they are connected by syndicate or alliance ties. Importantly, however, alliances are primarily associations focused on conflict resolution, while decisions concerning the internal discipline or military efforts remain entirely at the discretion of the single bosses – unless they directly impinge on other federation members (Hill, 2003, p. 72). Thus, syndicates are meant to represent the highest organizational level within the yakuza system. As such, we propose:

> H2a. Yakuza groups are less likely to attack other groups from other syndicates with which their syndicate has an alliance.
> H2b. Yakuza groups are less likely to retaliate against other groups from other syndicates with which their syndicate has an alliance.
> H2c. Yakuza groups are less likely to retaliate on behalf of other groups from other syndicates with which their syndicate has an alliance.

These higher-level coordination channels – both within and between syndicates – both enable dispute mediation and embed groups within a multilevel structure. Syndicate structures and alliances provide a means for groups to non-violently resolve disputes before they escalate to the level of violence and to stop an initial attack from leading to cascading cycles of violence. These structures also situate local yakuza groups within a larger organizational web that impacts their decision-making. As discussed in Section 2.2.1, street gangs often engage in violence based on status considerations – not retaliating or reacting after being attacked could cause them to 'lose face' (Gravel et al., 2018; Lewis & Papachristos, 2020). Yakuza groups face similar pressures; however, they must also consider how their actions can have implications for the rest of their syndicate and their syndicate's allies. Further, syndicates and alliances provide mechanisms like expulsion that increase the cost of retaliatory violence and other actions that could lead to complex status-oriented patterns of violence. As such, we expect:

> H3. The conflict network exhibits low levels of status-oriented conflict structures[7] seen in the street gang literature:
>
> > H3a. Yakuza groups are less likely to directly retaliate against another group that attacks them, regardless of the original attacker's syndicate affiliation.
> > H3b: Yakuza groups are less likely to form two-paths where they attack a third group after being attacked.
> > H3c: Yakuza groups are less likely to form out-stars where one group attacks multiple other groups.

[7] See Appendix A for visualizations of these structures.

> H3d: Yakuza groups are less likely to form in-stars where multiple groups attack one group.

Despite these multilevel considerations, yakuza groups still retain some agency when deciding whether and which group to attack. We expect this to result in yakuza groups attacking other groups with whom their territory overlaps or borders, similar to prior findings on street gangs in the existing literature. Territorial proximity provides opportunities for groups to attack one another and can generate motivations for conflict, such as competition for territory or business opportunities. Thus, we propose:

H4. Yakuza groups are more likely to attack groups headquartered:
- H4a. in the same city
- H4b. in an adjacent city
- H4c: in the same prefecture

2.3 Data and Methods

2.3.1 Data Overview

To investigate the patterns of violence in yakuza conflicts, we have collected a database of violent incidents between local yakuza groups across Japan from publicly-available media sources for the six-year period from 2014 to 2019. Before 2014, conflicts were so rare that a yakuza magazine published a special report asking "Why are there no conflicts?" (JJ11/2014). In 2015, a schism within the Rokudaime Yamaguchi-gumi caused an increase in violence: the rising number of incidents allowed us to collect enough data to produce reliable results. Recorded incidents include both attempted and executed violent acts – e.g., both attempted shootings and actual shootings were recorded – and range from homicides with firearms and knives to fistfights, hit-and-run car attacks, and property damage. While the North American-focused literature often concentrates on a narrower range of incident types – typically either gun homicides or both homicides and attempted homicides – the wider set of actions used here was chosen because these other types of incidents are also indicative of the conflict that we are investigating. Further, Japan's strict gun control laws make gun homicides exceedingly rare in Japan, meaning that violent attacks must be done by other means. Only incidents involving attempted or actual damage to people or property were included. Less severe incidents were not included; for example, an instance in which two yakuza members shouted at each other in a bar would not be included, but the incident would rise to the threshold of inclusion if the yakuza members then began fighting or started a brawl. Overall, the incident database includes 122[8] incidents between 170 local yakuza groups, with groups located in 31 Japanese prefectures and affiliated with 10 different yakuza syndicates.

The incident database and further information were used to create a multilevel network. While single-level networks include a set of nodes that represent actors or groups and a set of edges that describe a certain relationship between these nodes – like friendship or kinship ties – a multilevel network includes multiple sets of nodes of distinct types that are separated into different levels and multiple sets of edges that represent relationships between nodes in the same level and between nodes across levels (Lomi et al., 2016; Wang et al., 2013). Multilevel networks have been used to describe both legal-market and criminal structures (Coutinho et al., 2020; Hollway & Koskinen, 2016; Lazega et al., 2008; Stys et al., 2020). The multilevel network based on the incident database includes three levels of nodes and five sets of edges, summarized in Table 2.1. The network consists of three 1-mode networks representing local group violent conflict, syndicate alliances, and the geographic adjacency of cities and two

[8] Note that three incidents between members of the same group have been excluded and attacks that included multiple attackers or victims have been counted as one incident for each attacker-victim pair.

Table 2.1 Yakuza Conflict Network Level Information

	Conflict	*Alliance*	*Adjacency*	*Affiliation*	*HQ Location*
Node Types	Groups	Syndicates	Cities	Groups and Syndicates	Groups and Cities
Edge Type	Directed	Undirected	Undirected	Undirected	Undirected
Network Type	Unipartite	Unipartite	Unipartite	Bipartite	Bipartite
Nodes	170	10	79	180	249
Edges	114	12	32	170	170

bipartite networks representing affiliations of local groups to syndicates and the location of local group headquarters in different cities. Note that each local group is affiliated with one syndicate and headquarters are located in one city. The network representing violent conflict between local groups has been binarized, with an edge representing at least one violent attack between two groups; this was done because most edges (94%) had an edge weight of 1, while few edges had weights of 2 (5%) or 3 (1%).

The final level of nodes – representing which cities are within which prefectures and which prefectures are geographically adjacent to each other – is instead represented as structural zeros. The incident database includes groups from 31 of Japan's 47 prefectures, meaning that the resulting multilevel network is national in scope. Our analyses make use of simulation-based approaches, requiring that the simulated networks be similar to the observed network. To ensure that simulated networks would be comparable to the observed data, we restricted the simulation-based analyses and the multilevel ERGM by using structural zeros to constrain the simulations. Structural zeros fix the tie weight between two nodes at zero and do not allow this weight to vary during simulations; in effect, a given pair of nodes with a structural zero is not allowed to form a tie during any simulations. Given the wide geographic span of the multilevel network, we constrain network simulation to only allow attack ties to form between spatially proximate groups. Following exploratory data analysis, we determined that attacks occurred between groups that were either headquartered in the same prefecture, within adjacent prefectures, or within non-adjacent prefectures in a limited number of cases. This was used to create a target prefecture set for each prefecture – for a given prefecture, its target prefecture set includes itself and any adjacent prefectures, as well as any non-adjacent prefectures with an attack in the incident database involving the given prefecture; this occurred for 11 pairs of non-adjacent prefectures, typically prefectures that are close to one another but are separated by a small section of another prefecture. In the observed network, no attack ties exist between groups that are not within each other's target prefecture set. Using structural zeros, we fixed this structure within the network – for example, Tokyo and Hokkaido are non-adjacent prefectures and the incident database does not include any attacks between groups from these prefectures; thus, we constrained the network such that no ties can exist between a group from Tokyo and a group from Hokkaido in any simulations. On the other hand, we allowed variation for ties between a group headquartered in Tokyo and all other groups headquartered in a prefecture within Tokyo's target prefecture set. When structural zeros are considered, the group-to-group network density of observed edges as a proportion of the total possible edges is 0.015.

Further, we used qualitative interview findings to contextualize these quantitative data. Semi-structured interviews have been conducted with four hard-to-reach informants who include the boss of an influential yakuza group, the ex-legal counsel of one of the major yakuza groups, the chief of a prefectural police force, and a journalist specialized in yakuza affairs. Interviewees have been selected by the authors for their direct knowledge of the topic and have been reached through the network of one of the authors. All interviews have been conducted in person in Japan between April and July 2023 and have been anonymized.

2.3.2 Data Collection and Limitations

Data on incidents were collected from two primary types of sources: (1) major national newspapers like the Yomiuri Shinbun and Asahi Shinbun and (2) yakuza-focused news aggregators like Yakuza893 and YakuzaNews and general news aggregators like Factiva that link to national, regional, and local news sources. Both sets of sources were used in conjunction to ensure proper coverage of all relevant events and to ensure validity by triangulating incident details and confirming consistency across sources. Furthermore, research on the yakuza has the advantage of allowing researchers to use two additional information sources. First, we used yakuza-focused weekly and monthly magazines, most notably the *Jitsuwa Jidai*, to collect information on the alliances, conflicts, and meetings among yakuza syndicates. These magazines have been used as part of prior research (Hill, 2003; Kaplan & Dubro, 2012) and have been described as the “de-facto trade periodicals” for the yakuza (Adelstein, 2009). Second, we have used Yakuza Wiki as a source of information on the location of the headquarter offices of local yakuza groups. While the contributors to the website are anonymous, the information can be considered reliable and updated: doublechecks with other sources have invariably confirmed the information provided by the website and journalists with a long career of reporting on the yakuza have confirmed the validity of the site in interviews. On one occasion, the two authors visited the location of five yakuza offices as indicated on the website and found that they were accurate. This media site is similar to PrimeCrime.ru, another site used to study the Russian Mafia (Catino, 2019; Tsereteli, 2022; Varese, 2017; Varese et al., 2021).

Potential limitations to this study are primarily related to the nature of the data and the sources from which it was collected. The use of news media introduces the potential for reporting bias, where newspapers may not have reported on all relevant incidents during the investigation period. Incidents considered too insignificant may have been observed by newspapers but not reported, while certain incidents may not have been observed at all; for example, a murder in which the body has been concealed would likely be reported as a missing persons case rather than a violent incident. This limitation is mitigated by the fact that we collected incidents from a variety of sources that ranged in focus from national- to local-level. This broad coverage – especially the use of news aggregators pulling from local news sources that are more likely to contain reporting on lower-level incidents that might not appear in national news sources – decreases the likelihood of missed incidents. Further, interviews with the police suggest that our approach captured most of the relevant incidents from the investigation period, as the majority of incidents were reported in the local news (Interview S2).

Another limitation is that our data only includes yakuza groups that were involved in at least one incident as either attacker or victim; as such, there are yakuza groups that are not included as nodes in the network. This exclusion of yakuza groups that were not involved in any attacks – nodes that would be isolates if included – impacts the density of the group-to-group attack network, as well as its in-degree and out-degree distributions by decreasing the number of nodes that have in-degrees and out-degrees of 0. Any potential limitations stemming from this boundary specification decision would be related to the multilevel ERGM, as any nodes not involved in an incident are not part of the incident database; thus, there would not be a material impact on the temporal and atemporal multilevel reciprocity measures. This limitation is mitigated by the fact that our focus is primarily on what a yakuza group does after an initial attack, as this is most revealing about the organizational structure of the yakuza and how this structure impacts decisions around violence. However, this does suggest some caution must be taken when interpreting ERGM parameters relating to the likelihood of an initial attack, like terms investigating the overall likelihood of attacks, the effect of belonging to a specific syndicate, or the effect of territory overlap or adjacency. For example, the inclusion of these isolated nodes would likely decrease the edges term – which captures the baseline likelihood of tie formation in the model – and the syndicate activity control variables and could potentially decrease the magnitude of positive territory overlap and adjacency coefficients. While some caution is warranted, we have elected to keep these parameters in the final model since these factors have been suggested to be important based on prior findings in the literature and their interpretation can still be helpful in testing our hypotheses.

Another potential limitation is that the conflict network is sparse, with a density of 0.015 and average in-degree and out-degree of 0.67. However, street gang conflict networks also tend to be sparse, as noted previously in the existing literature (Bichler et al., 2020; Lewis & Papachristos, 2020; Randle & Bichler, 2017). Prior research on mafias and organized crime groups also suggests that conflict networks involving these groups are likely to be even sparser than equivalent networks for street gangs (Catino, 2014). Our network is in-line with prior research and theoretical expectations for density and connectivity. Finally, two incidents included multiple potential attacking groups due to vague reporting. The analyses in Section 2.4 were run using alternate attacking groups and did not show a material change; these analyses are available from the authors upon request.

2.3.3 Analytical Approach

To investigate our hypotheses, our analysis proceeded in two phases. First, we defined new temporal and atemporal measures for use with the incident database, taking incidents as the unit of analysis. The statistical significance of these results was then tested using a simulation-based approach similar to Conditional Uniform Graph (CUG) tests adapted for temporal networks. Second, we used a multilevel exponential random graph model (ERGM) to model the tie formation process and test hypotheses about the mechanisms that impacted groups' choices to attack one another.

The first phase of analysis focused on the temporal view provided by the incident database. Similar to the local group-level network described in Table 2.1, the incident database provides a network view of violence between local yakuza groups. While the local group-level network – like most gang violence networks analyzed in the existing literature – provides a static view where all edges are included regardless of when they occurred, the incident database includes information about the date on which a given incident occurred. This allows us to better understand the results by considering the timeframe in which they occurred. We have defined new temporal multilevel measures of retaliation, illustrated below in Figure 2.2a-d. Simple retaliation refers to the original victim retaliating directly against their attacker. Direct syndicate retaliation refers to a case where another group in the same syndicate as the initial victim directly retaliates on their behalf. Indirect syndicate retaliation, on the other hand, represents the case where retaliation occurs purely between syndicates rather than local groups. Finally, allied retaliation represents the case where groups whose syndicates are allied directly retaliate for each other. Each of these measures accounts for the temporal nature of the data, where the result ranges between 0 and 1 and represents the percentage of incidents where at least one instance of the corresponding kind of retaliatory attack occurred within a certain time period following the initial incident. For example, a value of 0.5 for direct syndicate retaliation within seven days would suggest that in half of the cases where A attacks B, a subsequent attack occurs from B's syndicate-mate C to A within seven days of the original A-B attack. Note that the time period considered begins on the first day after the incident considered, as the reporting of incidents typically did not include the time at which they occurred. We also applied two atemporal measures to the incident database: (1) the share of incidents that occurred between groups affiliated with the same syndicate and (2) the share of incidents that

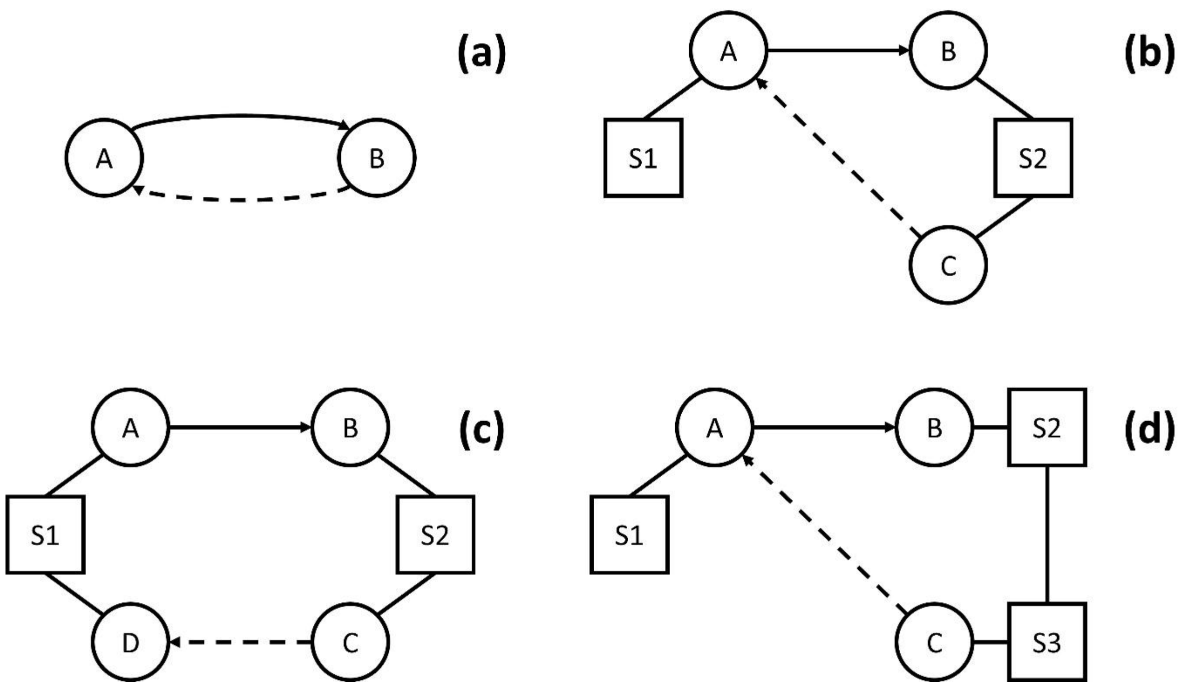


Figure 2.2. Multilevel network configurations for (a) simple retaliation, (b) direct syndicate retaliation, (c) indirect syndicate retaliation, and (d) allied retaliation. Circular nodes represent local yakuza groups, while square nodes represent syndicates; solid lines represent edges that already exist or have already occurred, while dashed lines represent the edge of interest occurring after the solid-line edges.

occurred between groups affiliated with syndicates that are allied.

To assess the statistical significance of the results, we have adapted the Conditional Uniform Graph (CUG) test methodology in which an observed value of a measure is compared to the values when the measure is taken for a range of random networks simulated based on an appropriate null-model; the observed value is compared to the sample of simulated values to determine if the observed value is more extreme than what would be expected at random to generate p-values (Butts, 2008). Our approach simulates random incident databases by first placing a restriction on which groups out of those observed in the incident database are potential targets for a given attacking group. This restriction is based on the structural zeros described in Section 2.3.1 – a given attacking group is allowed to target groups that are in its prefecture's target prefecture group. The null-model used is conditioned on the out-degree distribution and the timing of attacks in the observed incident database. For each incident, the attacking group and the date of the attack are kept the same; only the victim group is varied within the simulations. The new victim group is chosen at random from the set of potential targets available to the attacking group based on the geographic restrictions, including the observed victim. This approach controls for network density, as the number and timing of attacks do not vary. Past network configurations do not influence the probability of a given victim group being chosen. 5,000 random networks are simulated for each measure to generate empirical p-values; these empirical p-values represent the proportion of simulations for which the test statistic is *equal to or more extreme* than the observed statistic for the incident database (Stadtfeld & Amati, 2021).

The second phase consists of an atemporal analysis that investigates the social mechanisms underlying the tie formation process in the yakuza conflict network using a multilevel exponential random graph model (ERGM). ERGMs allow for hypothesis testing for networks even though network data inherently violates the assumption of independence of observations required for traditional statistical models. ERGMs use Markov Chain Monte Carlo maximum likelihood estimation to estimate parameters and simulate networks to generate confidence intervals (see Hunter, Goodreau, et al., 2008; Hunter, Handcock, et al., 2008; Lusher et al., 2012; Robins et al., 2007 for technical discussion). These models include parameters for micro-level local configurations in the network that are theoretically tied to social mechanisms in the tie formation process; global network structures then arise from observed local configurations. Multilevel ERGMs extend the ERGM framework to include dependencies both within and across levels (Wang et al., 2013). In the multilevel ERGM results presented in Section 2.4.3, alliance ties between syndicates, affiliation ties of groups to syndicates, affiliation ties of groups to cities, geographic adjacency ties between cities, and affiliation ties of cities to prefectures are fixed as our hypotheses focus on attacks between yakuza groups conditional on these other ties; only group-to-group ties are allowed to vary during the model estimation process. Since all ties outside of the group-to-group conflict layer are fixed, each group is only affiliated to one syndicate and one city and each city is only affiliated to one prefecture, and the related ERGM parameters are included to test the effect of shared affiliations on tie formation, we are able to operationalize these affiliation ties between levels as node attributes. Parameters for multilevel dependencies – those dependent upon alliance ties between syndicates and geographic adjacency ties between cities – are instead operationalized as dyadic covariates similar to the treatment in Coutinho et al. (2020) and Stys et al. (2020). These effects cannot be captured by node attributes and instead are better represented by dyadic covariates that represent, for each pair of groups, whether the two syndicates to which they are affiliated are allied and whether the two cities in which they are headquartered are geographically adjacent.

To test the hypotheses laid out in Section 2.2.2, we included a number of single-level and multilevel parameters during the model specification process. To test H1a and H2a, we include effects for whether two groups are affiliated to the same syndicate and whether two groups are part of syndicates that share an alliance tie. For H3, we include parameters to assess the likelihood of status-oriented local configurations like retaliatory violence, two-paths, and out- and in-stars, which are modeled as geometrically weighted out-degree (GWODEGREE) and geometrically weighted in-degree (GWIDEGREE) respectively. GWODEGREE and GWIDEGREE represent out-degree centralization and in-degree centralization, where negative coefficients suggest centralization and positive coefficients suggest decentralization (Levy, 2016a, 2016b). Note that, as seen below in Section 2.4.1, the lack of any transitive triads or generalized exchange configurations in the observed network mean that these parameters cannot be included in the ERGM. Further, we include interaction terms for the likelihood of retaliation between groups affiliated with the same syndicate and between groups whose syndicates are allied to assess whether these structures are able to decrease the likelihood of retaliatory violence when controlling for the likelihood that an initial attack will take place; these parameters relate to H1b and H2b. For H4a-c, we include effects for whether two groups are headquartered in the same city, in adjacent cities, or in the same prefecture. We control for syndicate activity – whether groups affiliated with a given syndicate are likely to attack or be attacked more often than groups from other syndicates. Finally, structural zeros were used to place the same geographic constraint on edges as used in the simulation test procedure used in the first phase of the analysis – this constraint ensures that no ties are formed from an attacking group to a victim group not within the attacking group's target prefecture set.

2.4 Results

2.4.1 Descriptive Results

The visualization of the single-level group-to-group conflict network is presented in Figure 2.3. The network only contains 10 of the 25 yakuza syndicates and only 170 affiliated groups of thousands of such groups throughout Japan (NPA, 2022). Only a fraction of groups and syndicates were involved in violent conflict as attackers or victims during the investigation period – the majority did not engage in conflict on either side (JJ11/2014; JJ02/2015). As noted in Section 2.3.1, the conflict network has a density of 0.015 when accounting for structural zeros. Compared to the more cohesive Chicago gang homicide network from Lewis and Papachristos (2020) in Figure 2.3b, the visualization in Figure 2.3a shows that, when not considering multilevel dynamics, the group-to-group conflict network is primarily made up of simple network structures like lone or reciprocated edges between individual dyads that are disconnected from other nodes. This is confirmed using a triad census, with results for relevant triadic structures presented in Appendix B: the conflict network exhibits no local configurations for transitive triads or generalized exchange and few instances of other configurations theorized in the literature to represent more complex status-oriented violence like two-paths and out- and in-stars. Instead, the triad census is dominated by single edges from i to j or reciprocated edges between i and j. These descriptive results suggest that yakuza groups in Japan experience generally low levels of conflict and that when violence does occur, it rarely leads to repeated violence and complex network structures.

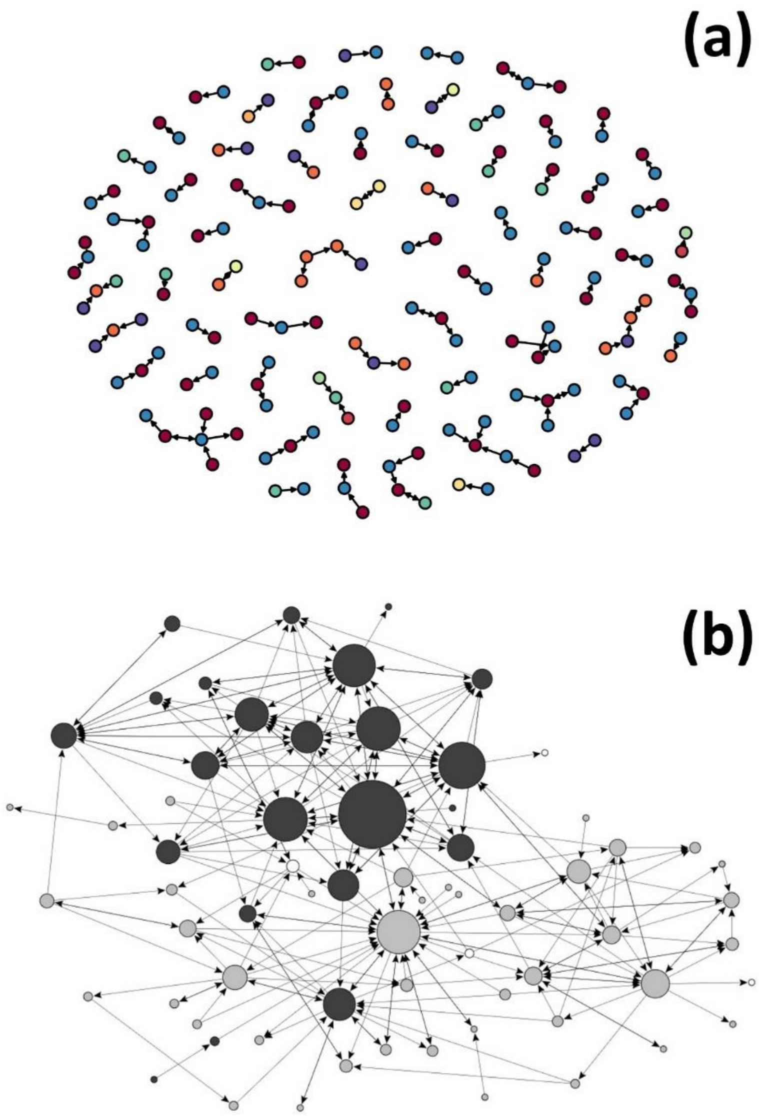


Figure 2.3. Conflict network visualizations. (a) Group-on-group yakuza conflict network; node colors represent the syndicate to which a group is affiliated. (b) Gang homicide network in Chicago from 1996-2000 from Lewis and Papachristos *(2020)*; node size is proportionate to the number of neighborhoods in which the gang was involved in a homicide, with node colors representing Black gangs (dark grey nodes), Latino gangs (light grey nodes), and white gangs (unshaded nodes).[9]

[9] Figure 2.3b reproduced from Lewis, K., & Papachristos, A. V. (2020). Rules of the Game: Exponential Random Graph Models of a Gang Homicide Network. *Social Forces*, *98*(40), 1829–1858. https://doi.org/10.1093/sf/soz106. Reproduced with the permission of Oxford University Press and the authors.

2.4.2 Simulation Test Results

Using the measures and simulation test procedure outlined in Section 2.3.3, we find preliminary support for hypotheses H1a, H1c, H2a, and H2c. Results for the atemporal and temporal measures are presented in Tables 2.2 and 2.3, respectively. The atemporal measures include metrics for the share of attacks between groups from the same syndicate, the share of attacks between groups from allied syndicates, and a single-level atemporal measure of retaliation typically used for similar networks elsewhere in the literature. 10% of attacks in the incident database were intra-syndicate, a lower rate than expected at random ($p_{lower} = 0$). 16% of attacks were intra-alliance, although this result was not statistically significant. The atemporal reciprocity measure applied to the conflict network shows that 21% of edges are reciprocated, a higher rate than expected at random ($p_{higher} = 0$).

Temporal reciprocity measure results are presented across a range of timespans following a given incident from a single day to the entirety of the rest of the period. Simple retaliation rates are low – only 6% of incidents are retaliated in this way even with the maximal timeframe – but higher than expected at random ($p_{higher} < 0.001$ for all timeframes).[10] Direct syndicate retaliation shows similar results, although the measure does not become statistically significant until a timeframe of at least 30 days is considered. Indirect syndicate retaliation shows high values across the range of timeframes shown, with statistically significant values for all timeframes except for $t = 1$. Within the first week following an attack, 25% of incidents are retaliated by another group from the same syndicate as the victim against a group from the same syndicate as the attacker; this rises to 81% of incidents when considering the entirety of the rest of the period. The results for longer timeframes need to be interpreted with some

[10] A further 4% of incidents were retaliated in this way on the same day as the initial incident.

Table 2.2 Atemporal Measures – Conditional Uniform Graph Test Results

	Rate	p_{higher}	p_{lower}
Intra-Syndicate Attacks	0.10	1.00	0.00
Intra-Alliance Attacks	0.16	0.17	0.90
Network Reciprocity	0.21	0.00	1.00
Empirical p-values based on 5,000 simulations			

Table 2.3 Temporal Measures – Conditional Uniform Graph Test Results

Timeframe (days)	Simple Reciprocity	Direct Syndicate Reciprocity	Indirect Syndicate Reciprocity	Allied Reciprocity
1	0.02 (0.00)	0.01 (0.22)	0.11 (0.60)	0.00 (1.00)
7	0.02 (0.00)	0.01 (0.23)	0.25 (0.00)	0.00 (1.00)
14	0.03 (0.00)	0.01 (0.22)	0.29 (0.00)	0.00 (1.00)
30	0.03 (0.00)	0.02 (0.02)	0.43 (0.00)	0.00 (1.00)
90	0.03 (0.00)	0.02 (<0.01)	0.61 (0.00)	0.00 (1.00)
180	0.03 (0.00)	0.02 (<0.01)	0.66 (0.00)	0.00 (1.00)
365	0.06 (0.00)	0.04 (0.00)	0.67 (0.00)	0.01 (0.10)
730	0.06 (0.00)	0.05 (0.00)	0.76 (0.00)	0.01 (0.10)
Rest of Period	0.06 (0.00)	0.05 (0.00)	0.81 (0.00)	0.01 (0.10)

Empirical p-values in parentheses based on 5,000 simulations; all p-value indicators for p_{higher}

caution. While yakuza groups possess an organizational memory like street gangs (Papachristos, 2009) and an initial attack can lead to a lasting enmity between groups or syndicates, a long period between incidents reduces the likelihood that two attacks are directly related. Further, these results are likely partially inflated due to the high rate of attacks between groups affiliated with the Rokudaime Yamaguchi-gumi, the Kōbe Yamaguchi-gumi, and Kizuna-kai. Despite these considerations, the high values even for shorter timeframes suggest that there is a genuine tendency towards this form of retaliation between syndicates. Finally, allied retaliation rates are low – only 1 incident (0.8%) was directly retaliated by an allied group more than 6 months following the initial attack – and are not significantly different than the values expected at random.

2.4.3 Multilevel ERGM Results

To further test our hypotheses, we specified a multilevel ERGM including the terms laid out in Section 2.3.3. Model results are shown in Table 2.4. The edges term – capturing the baseline likelihood of tie formation and interpretable similar to the intercept in regression models – is negative and significant, suggesting a generally low likelihood of tie formation in-line with the observed low density of the group-to-group network. The negative and significant Same Syndicate and Allied Syndicate terms suggest that a yakuza group is less likely to attack another group that is either part of the same syndicate or part of a syndicate to which its syndicate is allied, supporting H1a and H2a. This effect is stronger for shared syndicate affiliation than allied syndicate affiliation. We further find support for H4a and H4c, with the Same City HQ and Same Prefecture HQ terms being positive and significant – these suggest that a yakuza group is more likely to attack another group that is headquartered in the same city and the same prefecture. We do not find a significant effect for yakuza groups being

Table 2.4 Multilevel Exponential Random Graph Model Results

Parameter	*Estimate*	*SE*
Edges	-6.70***	2.10
Attributes		
Same Syndicate	-3.06***	0.57
Allied Syndicates	-1.84***	0.60
Same City HQ	1.83***	0.34
Adjacent Cities HQ	0.61	0.42
Same Prefecture HQ	1.04***	0.30
Structure		
Retaliation	-2.67***	0.74
Retaliation - Same Syndicate Interaction	2.44**	1.23
Retaliation - Allied Syndicate Interaction	0.12	1.01
Two-paths	-2.65***	0.36
GWODEGREE	6.28***	0.88
GWIDEGREE	3.81***	0.67
Controls		
Syndicate 1 Activity (base)		
Syndicate 2 Activity	-2.46	1.67
Syndicate 3 Activity	-0.30	1.01
Syndicate 4 Activity	-3.50*	2.11
Syndicate 5 Activity	-1.60	1.07
Syndicate 6 Activity	0.75	1.50
Syndicate 7 Activity	-2.55	1.63
Syndicate 8 Activity	-2.86**	1.15
Syndicate 9 Activity	-1.07	0.99
Syndicate 10 Activity	-1.06	1.07

***p<0.01, **p<0.05, *p<0.1; alpha = 0.5 for geometrically weighted terms headquartered in geographically adjacent cities when the other geographic parameters are included; as such, H4b is not supported.

The negative and significant Reciprocity term suggests that groups are less likely to engage in retaliatory violence than expected at random given the other parameters within the model and the model constraints, supporting H3a. We further find support for H1b and H2b. The interaction effect for retaliation between groups affiliated with the same syndicate – 'Reciprocity – Same Syndicate Interaction' – is positive and significant. Importantly, this term is a combination of the general retaliation effect and the same syndicate effect; when summed, the three terms add to create an effect size that is negative and of greater magnitude than the general retaliation term on its own, suggesting that retaliation is less likely between groups in the same syndicate. Similarly, while the interaction effect for retaliation between groups from allied syndicates is not statistically significant, the sum of the significant general retaliation and allied syndicate effects is negative and greater than the magnitude of the general retaliation term on its own; this suggests that groups belonging to allied syndicates are less likely to retaliate against one another given an initial intra-alliance attack occurs. The two-paths term is negative and significant, meaning that chain-like structures are less likely than expected at random. Finally, the geometrically weighted out-degree and in-degree terms are positive and significant, indicating degree decentralization (Levy, 2016a, 2016b). This suggests that out- and in-stars are less likely than expected at random given other effects included in the model and the model constraints, meaning that there are not groups that either launch attacks at or are victimized by a large number of other groups. These findings support H3b-d. Finally, two of our control variables for syndicate activity are negative and significant, suggesting that groups from the Kanto Sekine-ikka (Syndicate 4) and Kizuna-kai (Syndicate 8) are less likely to be

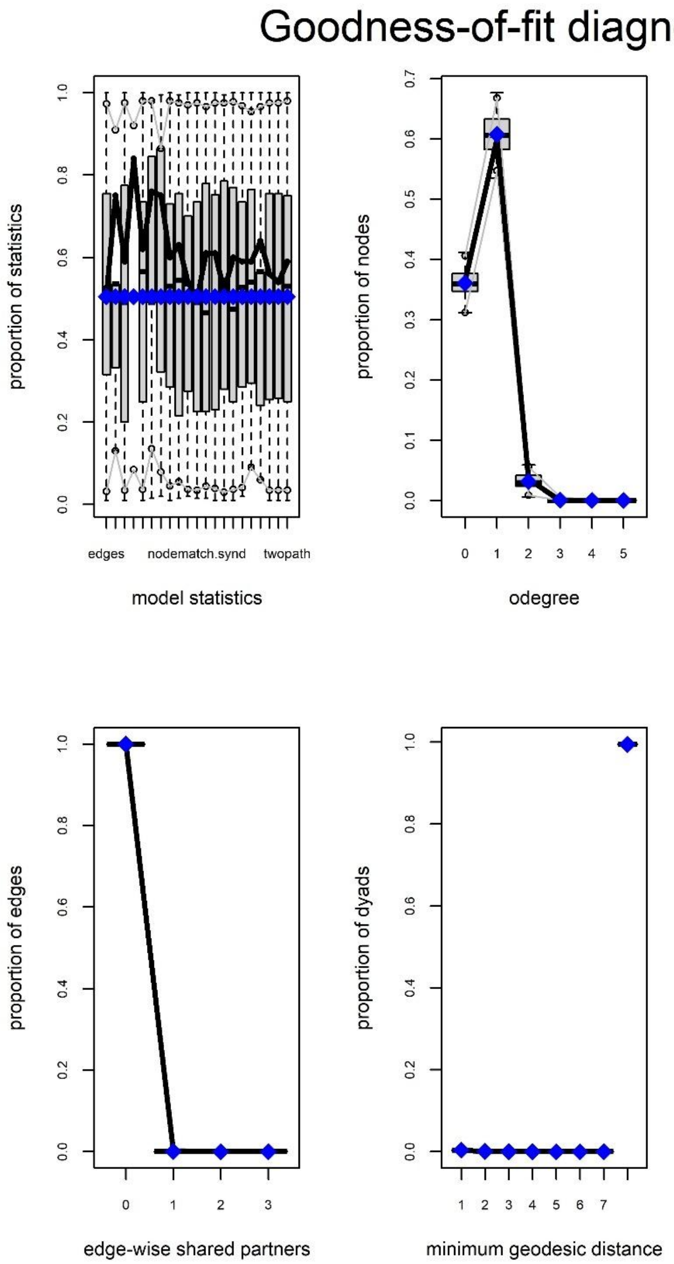


Figure 2.4. Goodness-of-fit plots for the multilevel exponential random graph model for parameters included in the model ('model statistics'), out-degree ('odegree'), in-degree ('idegree'), edgewise shared partners, and minimum geodesic distance.

involved in attacks as either the attacker or the victim than groups from Aizukosetsu-kai, the syndicate used as the reference category. These controls are not part of our hypotheses and these syndicates have few affiliated groups within the group-to-group attack network, limiting interpretability.

The model was specified in R using the *ergm* package (Hunter, Handcock, et al., 2008). Goodness-of-fit indicators are presented in Figure 2.4 and suggest that the specified model tended to capture the structural characteristics of the observed network that were not explicitly modelled (Hunter, Handcock, et al., 2008).

2.5 Discussion

The results presented in the prior section support the hypotheses laid out in Section 2.2.2. Yakuza syndicates function as cohesive organizations with a shared identity and the ability to constrain constituent groups' actions, with affiliated groups both avoiding conflict amongst themselves and retaliating on each other's behalf. Crucially, we find that this retaliation typically takes the form of attacking another group affiliated with an opponent's syndicate rather than the opponent group itself, further supporting the claim that yakuza groups identify themselves and other groups with their associated syndicates. Alliances, on the other hand, do not create a sense of shared identity between the groups of member syndicates: while alliances are successful at reducing conflict internally, groups do not retaliate on behalf of their allies. The conflict network was primarily composed of simple network structures that are not indicative of status-oriented violence, with lower rates of retaliatory violence, two-paths, and out- and in-stars than expected at random when controlling for other variables. The observed network also does not include any instances of complex configurations like transitive triads or generalized exchange. Finally, local market conditions and spatial proximity increased the likelihood of a group attacking another group that was headquartered in the same city or an adjacent city. Thus, we find that although yakuza groups do exist within national syndicate structures, an initial attack is still driven by local factors.

These results are supported by our interview findings and prior research. While the generally low rates of crime in Japan, especially violent crime, may be partly responsible for the low number of incidents recorded, there are also material conditions specific to the yakuza that contribute to this outcome. For the yakuza, like all organized crime groups, violence is costly. Japan's Firearm and Sword Possession Control Law significantly restricts the acquisition of firearms and imposes heavy penalties on criminal offenders if they use a weapon in the course of their crimes. While these features, among others, partially determine low levels of violent crime in Japan overall, other factors further increase the cost of violence for yakuza groups specifically. Attacking another group requires planning, resources, and labor, more so than most instances of individual civilian violence. It also carries the risk of death at the hands of an opponent or arrest by the police, carrying an opportunity cost as incapacitated yakuza members cannot generate revenue for their group. Further, yakuza groups – as with many mafias globally – financially support incarcerated members and their families. One yakuza boss in an interview suggested that, overall, a conflict could cost 500 million yen in total – approximately £3 million (Interview S1). Not only is the individual perpetrator liable for arrest, but an amendment to the 1992 Anti-Yakuza Law introduced in 2004 established the liability of gangs' representatives for acts of violence committed by members in conflicts (NPA, 2012), similar to the Racketeer Influenced and Corrupt Organizations (RICO) act in the United States. The pressures put on syndicates by the liability for the actions of affiliates and the stricter anti-yakuza regulations led to the current situation of low-intensity warfare that is taking place between the Rokudaime Yamaguchi-gumi and the Kōbe Yamaguchi-gumi (JJ04/2016). A full-blown conflict such as the Yama-ichi kōsō[11] would mean a complete halt of the two groups' activities (JJ12/2015). The effect of these regulations is best seen in how the pattern of violence has changed drastically since the implementation of the 1992 Anti-Yakuza Law and how within larger society political violence has virtually disappeared since the 1980s (JJ02/2015). Inter-group violence also increases police scrutiny of the attacker and victim, impacting their ability to operate; for example, groups involved in an attack may have their offices closed by police (Interview S4). Finally, yakuza groups must not only be concerned with their own interests, but also the interests of their wider syndicate. Even low-level conflicts could negatively affect the relationship between two national syndicates: "If two individuals belong to groups that are allied at the national level, and they know that, they will avoid big fights because it will create problems if someone high-ranking is involved" (Interview S3).

Furthermore, the yakuza is an oligopoly in which three syndicates – Rokudaime Yamaguchi-gumi, Inagawa-kai, and Sumiyoshi-kai – make up the majority of yakuza membership and the majority of group affiliations (NPA, 2022). Within the so-called 'Big Three', the Rokudaime Yamaguchi-gumi has held a virtual monopoly of force and had a balancing effect in the yakuza underworld. The current leadership of Tsukasa Shinobu has striven for stability and good relationships with all other groups, through participation in alliances and encouraging Yamaguchi-gumi group bosses to become "brothers" with external groups (JJ01/2008). Indeed, the Rokudaime Yamaguchi-gumi maintained the old alliances even after the internal split (JJ11/2015).

[11] The Yama-ichi kōsō (1985-1989) is a conflict involving the Yamaguchi-gumi and the Ichiwa-kai that resulted in 39 deaths.

Yakuza groups also have reasons to fight: they compete for economic opportunities and revenue streams, but do not compete in terms of the quality of the services they provide. Conquering a new territory or setting up a new market for one group often requires taking it from another group. The costs and opportunities for disputes incentivize yakuza groups, like all mafias, to avoid violent conflict and develop dispute resolution mechanisms that can either avoid conflict altogether or provide a quick end if it does occur, reducing the likelihood of drawn-out cyclical violence. Mafias around the world have developed higher-level coordination bodies to provide these mechanisms, like the regional and provincial commissions of the Sicilian Cosa Nostra and `Ndrangheta and the national commission and New York commission of the Italian-American Mafia (Catino, 2014, 2019). However, their need for secrecy often resulted in underspecified contracts and incomplete agreements, leading to misinterpretations and friction (Thompson, 2024). Yakuza groups, on the other hand, exist within a legal grey area that allows them to remain visible and make public commitments with less fear of police repression (Baradel, 2021). The reconciliation ceremony described in Section 2.2.2 is effective because it publicizes the cessation of conflict between two parties, the terms of the resolution, and how deviation from the agreement by either of the groups should be punished by third parties; these agreements are then further publicized by reports in magazines like the *Jitsuwa Jidai*. The terms of alliances between syndicates are similarly explicit and determined in advance (Mizoguchi & Suzuki, 2021). However, syndicates and alliances cannot solve internal problems and, indeed, in the past twenty years the major conflicts have been those internal to the syndicates and groups, such as conflict for leadership succession – such as the Dōjinkai split in 2007 (JJ01/2008) – or internal splits like the 2015 schism within the Yamaguchi-gumi (JJ10/2016).

One apparent contradiction between the network analysis results and interview findings concerns retaliatory violence. According to the former legal counselor for a major syndicate, the cultural norms of the yakuza require that attacks must be retaliated: "It is a rule of the yakuza: if you are attacked, you must strike back. Then you can proceed to the reconciliation" (Interview S3). This was echoed by a pair of journalists who explain that the groups that deviate from this rule are bound to be destroyed (Mizoguchi & Suzuki, 2021). At first glance, this seems to contradict the low rates of simple retaliation in the simulation test and ERGM results. However, this norm of retaliatory violence does not necessarily refer to retaliation between the original attacker and victim. Instead, the cohesive nature of yakuza syndicates means that attacks between other groups within the two syndicates – what we have called indirect syndicate retaliation – are sufficient for this norm. After an attack, both the attacking group and the victimized group are under greater police scrutiny since police know of this norm of retaliation, making it more difficult to plan and launch a retaliatory attack. Thus, other groups affiliated with the same syndicate can act on behalf of the victim, as seen in the simulation test results.

2.6 Conclusion

Prior research on violent conflict among street gangs and criminal groups has focused on how status-oriented violence leads to dense, complex webs of conflict where one attack begets another; even mafias that have developed higher-level coordination bodies between their operational units still transfer violence from one territory to another and face the risk of descending into cycles of retaliatory violence. The conflict resolution mechanisms provided by yakuza syndicates and alliances, on the other hand, result in markedly different patterns. While yakuza groups still initiate violence due to local market considerations similar to street gangs, an initial attack tends not to lead to direct retaliation or more complex network structures. Instead, groups are able to send issues or disputes up the chain of command to higher levels in their larger organizational structures for reconciliation in public settlements that increase groups' abilities to make credible commitments to avoid conflict. This process allows yakuza groups to avoid costly conflict. As Schelling (1980) notes, "Most conflict situations are essentially bargaining situations" (Schelling, 1980, p. 5). Yakuza syndicates made up of numerous affiliated local groups function as cohesive organizations: they are able to constrain the actions of their members and provide a collective identity that members seek to defend from attack. Alliances play a similar role in defusing conflict but fail to provide a collective identity for members – thus, syndicates operate as the top of the organizational structure in the yakuza.

Our findings contribute to the literature on the yakuza and on mafia structures, providing support for the argument that yakuza syndicates act as cohesive organizations rather than just loose brands or federations and provide mechanisms that can help reduce violent conflict. This contribution is enabled by our novel multilevel conceptualization of the conflict network and our novel network analysis measures. While only looking at the group-to-group conflict network fails to reveal complex attack structures, these attacks are embedded within a higher-level social system of syndicates and alliances that impacts the decision-making of local yakuza groups. Future research on street gang, mafia, and organized crime group conflict should account for these multilevel considerations when investigating why and how these groups fight. More generally, we contribute to the literature on how organized crime groups fight: while prior research has considered how a criminal group's structure impacts its use of violence, previous studies have not considered the dynamics of conflict as different groups attack back and forth.

Recent trends in the social, economic, and legal context of the yakuza suggest that the findings in this research may soon change. First, while the yakuza have traditionally enjoyed recognized status under Japanese law, groups have come under increasing law enforcement pressure over the past decade following the introduction of the Yakuza Exclusion Laws[12] in 2010. This pressure, combined with falling membership due to flagging recruitment rates and an aging member base, has led to economic challenges (Baradel, 2021), which could spur greater competition and conflict between groups fighting over a shrinking pie. Similarly, greater law enforcement pressure can spur conflict, as seen with the introduction of civil gang injunctions in Los Angeles (Bichler et al., 2019; Bichler et al., 2020, 2021). On the other hand, greater police scrutiny also increases the cost of violence and creates greater incentives to resolve disputes peacefully. These two issues pull the yakuza in opposite directions (Catino, 2019). Second, the yakuza's monopoly on organized crime in Japan has been under assault in recent years – in part due to the declining membership and economic troubles faced by the traditional syndicates – as other organized crime groups have moved into the Japanese market. While yakuza groups that are not affiliated with the same syndicate or part of an alliance structure can still make use of formalized conflict resolution mechanisms, although with greater difficulty, the introduction of new groups that do not share the yakuza's norms is likely to decrease the possibility of peaceful reconciliation. The structure of the yakuza often allows conflict situations to resolve as bargaining situations, but this transformation may not be possible between yakuza and non-yakuza groups, with the potential for rising violence as a

[12] *Bōryokudan Haijōjōrei*

result. The yakuza's structure is currently able to constrain groups and avoid complex and prolonged conflict patterns; however, these structural features may not be sufficient to limit violence in the coming years.

3

Factions and Brokers in the Russian Mafia: Investigating the Structure of the Thieves-in-Law

Abstract[13]

This paper investigates the internal structure of the fraternity of the *vory-v-zakone* – the high-ranking members of the Russian Mafia – using a co-signing network based on 50 signed edicts (*progony*) produced by the fraternity. We find that the *vory* are a multi-ethnic, polycentric association grouped into factions based on vertical patronage relationships; these factions are interconnected by key brokers that bridge between groupings. While this structure allows for effective collective action, it also points to possible fractures that can be exploited by authorities to weaken the fraternity. This research is the first quantitative network analysis of

[13] This work was done in collaboration with Prof. Federico Varese and Dr. Elena Racheva. A version of this paper has appeared in: Breuer, N., Varese, F., & Racheva, E. (2025). Factions and Brokers in the Russian Mafia: Investigating the Structure of the Thieves-in-Law. *The British Journal of Criminology*, azaf041. https://doi.org/10.1093/bjc/azaf041

documents produced by a higher-level coordination body within a mafia and contributes to the literature on the emergence of coalitions and factions within organizations.

3.1 Introduction

The *vory-v-zakone* (lit. 'thieves-in-law') are the members of a criminal fraternity that has dominated the prison system and wider criminal underworld in the Soviet and post-Soviet space for more than a century. All members, who share the same role and official authority, are prominent, respected bosses of criminal groups. Documented since the 1920s, the *vory* fraternity has survived despite external pressures like the dissolution of the Soviet Union and the increasing authoritarianism of Putin's regime, as well as internal threats like frequent internecine conflict. They have maintained organizational flexibility and adapted procedures to meet new challenges (Varese et al., 2021). This research aims to investigate the internal social structure of the *vory* and how bosses interact in the highest levels of a mafia.

A key coordination mechanism amongst the *vory* are formal, signed documents called *progony* that spell out rules and enforce decisions distributed across the post-Soviet criminal world. We collected an original database of these documents issued in 2013-2022 to construct co-signing networks. We then use descriptive and hypothesis-testing network analysis techniques to explore the underlying social structure represented by the decisions of *vory* to sign *progony* together. We find that the fraternity has developed a factionalized and polycentric organizational network structure with multiple centers of power that are interconnected by key brokers. While factions exist, clustering today is primarily driven by vertical patronage relationships rather than shared ethnicity or similar age, in contrast to narratives of ethnic and generational cleavages in the past among the *vory* in the existing literature.

This research contributes to the field's understanding of the structure of mafias by presenting, to our knowledge, the first direct analysis of the structure of higher-level coordination between bosses in a mafia. While prior research has furthered our understanding of numerous aspects of mafia structure, these analyses have typically focused on the lower-level structure of mafia families, the main operational units of the organizations. Higher-level coordination bodies that bring together the bosses of these operational units, on the other hand, have been more difficult to investigate because these associations lack written records and do not generate the same volume of evidence as the day-to-day operations of mafia families. Our novel database of *progony* allows us to directly scrutinize the internal proceedings of a mafia's governing body and the mechanisms that drive relationships between mafia bosses.

The rest of this paper is organized as follows. Section 3.2 presents a review of the existing literature on the *vory* and higher-level coordination bodies within mafias, as well as analytical expectations based on prior findings. Section 3.3 introduces the data and the methods. Section 3.4 presents the main results, while Section 3.5 discusses the findings. Section 3.6 concludes.

3.2 Literature Review and Expectations

3.2.1 *Vory* Background

The Russian Mafia is not a single criminal organization, nor is it exclusively Russian; instead, it is a constellation of affiliated groups that span the post-Soviet space with members from numerous countries and ethnic groups (Galeotti, 2004; Varese, 2001). These various organizations and associations – while distinct – are united by shared traditions and titles. Within this system, *vor-v-zakone*, often shortened to *vor*, is the title bestowed upon prominent, well-respected bosses in the post-Soviet prison system and wider criminal community; it is best rendered in English as 'criminal who adheres to the code' or 'thief-in-law' (Frisby, 1998; Serio & Razinkin, 1995). The term *vor-v-zakone* first emerged in the 1920s and 1930s to refer to professional criminals in the gulags and prison camps (Cheloukhine, 2008; Frisby, 1998). These criminals were respected by other members of the criminal community and lived by a code (*ponyatiya*), with rules prohibiting *vory* from officially marrying or cooperating with or working for the state and requiring *vory* to earn revenue solely from criminal activities (Gilinskiy & Kostjukovsky, 2004; Tsereteli, 2022).

For over a century, the *vory* have coordinated daily life in many post-Soviet prison systems – they collect money into and distribute from a common fund (*obshchak*) for the welfare of inmates, regulate interactions and transactions between prisoners, and resolve disputes; they also remain involved in external criminal activities (Cowley et al., 2015; Slade & Azbel, 2020). Since the 1990s, the *vory* have led many prominent criminal groups outside of prison that provide private protection and governance services to firms in both the illegal and legal markets, offering dispute resolution services and generally "[acting] as a vital lubricant for the cogs of organized crime activities" (Slade, 2014, p. 18). Today, the *vory* are largely concentrated in the post-Soviet states, particularly Russia and Ukraine; according to PrimeCrime.ru – a news agency focused on the *vory* – there are 399 active *vory*, of which 299 (75%) are at large (PrimeCrime.ru, n.d.).

Historically, the title of *vor* is only bestowed upon well-known criminals who are selected by existing *vory* for being respected as 'upstanding' members of the criminal community and abiding by the thieves' code. The candidate undergoes an observation period by the *vory* and their subordinates (Cheloukhine, 2008; Varese, 2001). During this time, the candidate is nominated by at least two existing members of the fraternity who act as sponsors and 'crowners'. After the observation period, a gathering of the *vory* (*skhodka*) is called; if the candidate has followed the thieves' code and deserves the title, he is crowned as a *vor* and officially admitted into the fraternity. The crowners act as guarantors of the newly-initiated member, vouching for their reputation (Cheloukhine, 2008, p. 359). Newly-crowned members often look to their crowners as patrons that provide protection and connections; in return, the crowners receive the support of their clients.

Once crowned, *vory* occupy top positions within hierarchical structures both within prisons and in organized crime groups. Prisons will typically have one key *vor* who then appoints a lieutenant (*polozhenets*) who manages daily life within the prison (Cheloukhine, 2008; Kupatadze, 2014). The lieutenant or the *vor* then appoints 'watchers' (*smatryashi*) who are in charge of aspects of prison life like administering the common fund (*obshchak*), regulating gambling, and managing communications within and between prisons (Kupatadze, 2014). Outside of prisons, the organized crime groups led by the *vory* take various forms, ranging from rigid hierarchies to loose associations of hierarchical subgroups (Catino, 2019; Galeotti, 2018).

The association of the *vory* is formally egalitarian – officially, all *vory* are equal and none hold a higher position or greater authority than another. Despite the lack of formal hierarchy within the association, however, some *vory* hold more power and influence (Cowley et al., 2015, p. 202). For example, Handelman (1995) reports a core of senior *vory* who made major decisions and an even smaller leadership core in the late 1980s and early 1990s (Handelman, 1995, p. 39). *Vory* frequently vie for power and seek to advance their personal interests.

These power struggles have been reported to lead to the emergence of factions that fragment the association as its members compete against each other (Slade, 2014, p. 106). Some prior studies suggest that these factions are based upon ethnic and generational cleavages, with *vory* of the same ethnicity and age partnering together. Mironova (2023) and Slade (2014) suggest that the *vory* began to ethnically segregate in the 1980s and 1990s respectively, citing ethnic divisions between the Georgian, Slavic, and Caucasian *vory* (Mironova, 2023, p. 69; Slade, 2014, p. 98). Other studies, however, challenge the idea that the *vory* are ethnically segregated. For example, Galeotti (2004) suggests that ethnic ties have decreased in importance within Russian organized crime generally (Galeotti, 2004, p. 58). Mironova (2023) and Slade (2014) also report that a generational divide emerged in the 1990s during the economic turbulence of the fall of the Soviet Union and the growth and expansion of the *vory*; the economic opportunity caused a split between older *vory* who were committed to the ideology and tradition of the *vory* and younger, newer *vory* who were focused on economic self-interest (Mironova, 2023, p. 166). Slade (2014) investigates a split within the Kutaisi-based Georgian *vory* in 2004, finding that the subgroup of older *vory* was centralized around a key leader, Tariel Oniani, and was densely interconnected; the younger *vory*, conversely, were decentralized, less interconnected, and lacked a clear leader (Slade, 2014, p. 106). Another proposed explanation for subgroup formation is vertical patronage relationships. Differences in power between the *vory* provide an incentive for weaker *vory* to align with and support more powerful members in return for protection (Slade, 2014). These alliances can then shift as the balance of power adjusts over time (Galeotti, 2004). Newly-crowned *vory* also often look to their crowners as patrons in return for economic or political support. These vertical patronage relationships can lead to informal authority accumulating within certain central *vory*, as those key players with many clients have power and authority that can attract new supporters.

The *vory* use two interrelated mechanisms for coordinating across their large and geographically-distributed fraternity. First, the *vory* hold meetings called *skhodki*. They can be convened by any *vor* (Slade, 2014, p. 97); any *vor* can attend, although typically only some do given the fraternity's geographic spread. Meetings occur within prisons or outside; further, there are locally- and regionally-focused *skhodki* as well as general meetings for all *vory*. *Skhodki* serve two main purposes. First, they act as criminal courts where *vory* can adjudicate disputes and, particularly in prisons, levy punishments for breaking the thieves' code (Cowley et al., 2015, p. 200; Slade, 2016; Slade & Azbel, 2020). Second, *skhodki* are the main deliberative meetings for the *vory*, providing an opportunity for the association to discuss issues facing the fraternity. Topics include strategies to respond to law enforcement activities, the creation of new rules and interpretation of existing rules, the administration of the common fund, the initiation of new *vory* and the expulsion of members deemed to have violated the thieves' code, and coordination between different *vory* and organized crime groups (Gilinskiy & Kostjukovsky, 2004). For example, in 2015 around 100 *vory* gathered for a *skhodka* in Armenia to resolve two major conflicts between criminal clans (PrimeCrime.ru, 2015). Similarly, in 2020 approximately 30 *vory* met in Istanbul, Turkey, to discuss the status of one of their members who had been expelled from the fraternity due to a conflict with another *vor* (PrimeCrime.ru, 2020).

While *skhodki* provide an opportunity for *vory* to deliberate, the second coordination mechanism – *progony* – provides a means of codifying, communicating, and coordinating the *vory*'s decisions. *Progony* are documents issued and signed by *vory*. Following a meeting, an edict is published to report the decisions made and to finalize the crowning of new *vory*; *progony* are also published without an affiliated meeting to propose courses of action, new rule interpretations, and the cancellation of the status of a *vor*. These documents represent the main steps of formalizing and legitimizing decisions within the association. While meetings are for deliberation, issuing a *progon* about the crowning of a new member, the stripping of the title from another, or the implementation of a new rule provides an opportunity for other *vory* – especially those who could not attend the meeting – to accept or challenge the ruling. This can result in a series of edicts back and forth between opposing *vory*. *Vory* sign *progony* either because they endorse their contents or to support a more powerful patron. While signatures are necessary for a *vor* to officially show they support a document, the geographic spread of the *vory* means that some may not be able to sign a *progon* in-person; in these cases, the *progon* can be sent to them by text message (Varese, 2017, 2024). To use an analogy to the U.S. legislative bodies, *skhodki* play a similar role to meetings and committees in the Congress – these meetings provide the opportunity to debate and vote on decisions. *Progony* are similar to bills and acts. A *progon* is issued and sponsored by its signatory *vory* proposing a new change or decision like a bill. Further *progony* are sometimes then issued that disagree with or support the original *progon*, similar to how legislators in a House of Representatives committee will write supporting and dissenting opinions about a bill before it is introduced to the wider body for debate and voting. *Skhodki* and *progony* function together to enable coordination amongst the *vory*. Meetings are used to discuss, while edicts are issued to allow input from other members, enable further deliberation and adjustments, and finally formalize and communicate the decision as official and enforceable.

3.2.2 Mafia Coordination

Similar to the *vory*, other mafias have developed higher-level coordination mechanisms for collective decision-making and communication between bosses and operational units. Historically, the Sicilian Cosa Nostra, the Calabrian `Ndrangheta, and the Italian-American Mafia established commissions or ‘cupola’ to enforce rules, resolve disputes within and between families, discuss common issues, and make organization-wide decisions (Gambetta, 1993; Varese, 2017). The Palermo Province Mafia Commission, formed in 1957, represented the 78 Cosa Nostra families in Palermo, while the Sicilian Regional Commission convened representatives from the 100-150 families on the island; each member of these bodies represented a district of three geographically-connected families (Catino, 2019, p. 156). This system endured until the 1990s, when it broke down due to increased state repression. In New York City, bosses of the city's five mafia families each had a seat on the New York City Commission (Jacobs, 2020; Thompson, 2024); a National Commission was also established in 1931 which fulfilled a similar function for families across the U.S., although it no longer operates (Catino, 2019; DellaPosta, 2017). Other mafias have developed similar coordination mechanisms. The 25 syndicates of the Japanese yakuza form alliance associations, like the Kantō Hatsuka-kai between the syndicates based in Tokyo and the wider Kantō region (Hill, 2003). The Triads in Hong Kong and Taiwan also had coordination bodies called headquarters or central committees, although these have largely become defunct (Catino, 2019; Chu, 2000; Zhang & Chin, 2003). These bodies and mechanisms enable bosses to make joint decisions, manage economic ventures, and resolve disputes to avoid conflict or law enforcement attention.

While all higher-level coordination bodies in mafias face coordination troubles as they try to remain secret to avoid drawing law enforcement attention, the *vory* face particular logistical challenges. Most coordination bodies historically have consisted of a small number of geographically concentrated families in places like Palermo and Sicily, New York City, and Hong Kong. While Japanese yakuza syndicates operate throughout Japan, organizing meetings is easier because syndicates are formally legal organizations and inter-syndicate associations are typically regionally-defined (Baradel, 2021). Further, coordination bodies in other mafias are often representative and only contain a subset of bosses; the *vory* fraternity, conversely, consists of all criminals who have gained the title. Finally, the *vory* come from a number of ethnic backgrounds and exhibit a wide range of ages, while other mafias are typically monoethnic and have bosses which tend to be older mafiosi (Gambetta, 1993; Musotto, 2022). Ethnicity and age present potential cleavages along which factions could form and divisions could arise, as has been reported for the *vory* in the past (Mironova, 2023; Slade, 2014).

The existing literature on the structure of mafias has largely focused on levels below the higher-level coordination bodies, such as the structure within and between families or operational units. Interactions between bosses within coordination bodies have historically been more difficult to investigate because these bodies produce less usable evidence than lower levels of the organizations. Coordination bodies meet infrequently and sporadically as needed to discuss important issues or settle disputes; these gatherings are nearly impossible for police to bug.[14] However, some studies have provided insight into the role of and interactions between bosses more generally. Much of the literature on the role of bosses in mafia structures has focused on the concept of brokerage. Brokers are actors that bridge structural holes between otherwise-disconnected people and groups, enabling the transmission of resources, information, and ideas within and between organizations (Burt, 1995). Brokers both benefit personally from their positioning and lead to beneficial outcomes for the wider network. For example, Burt (2004) examines the connections between supply chain-focused managers at a large electronics company. Connections tended to stay within business units, acting as

[14] An exception was the bugging of the meeting of the Lombardy Regional Commission of the `Ndrangheta during Operation *Infinito*. While the wider operation that contains this bugging has been used to investigate the structure of the `Ndrangheta (Calderoni, 2016; Calderoni et al., 2017; Calderoni & Superchi, 2019), these studies have focused on the organization's wider structure and the role of bosses as brokers rather than focusing on the relationships between different bosses from different families within the Commission.

organizationally-defined factions; connections between business units were enabled by a small number of brokers that formed ties across departments and synthesized information from multiple groups (Burt, 2004, p. 366). These gap-spanning managers enabled more effective coordination and information transmission. These brokers also had higher salaries, received better job evaluations, and were more likely to be promoted. Other studies suggest that this extends to the members of organized crime groups: criminal brokers are able to strategically position themselves to gain information, seize business opportunities, and insulate themselves from prosecution and law enforcement disruption (Morselli, 2001, 2003, 2010).

Prior research suggests that mafia bosses in particular act as brokers between both other mafiosi and between mafia families. In the early- and mid-20th century the Italian-American Mafia consisted of strongly intra-connected factions distributed across the U.S. that were interconnected by a small number of key brokers (DellaPosta, 2017, 2023). These actors who bridged between different families tended to be either high-status bosses or low-status outsiders who could not officially be initiated as mafiosi because they were not Italian. In particular, bosses tended to be generally well-connected both to their own family and between families, while the low-status outsiders bridged between groups despite generally being poorly connected to the network (DellaPosta, 2023). In his study on the criminal career of Sammy Gravano in the Italian-American Mafia in the 1970s and 1980s, Morselli (2003) finds that Gravano's brokerage positioning increased as he was promoted in the organization, gaining better access to business opportunities and decreasing his personal risk. Musotto (2022) investigates the structure of the Sicilian Cosa Nostra using data from a 2008 anti-mafia operation that thwarted an attempt to rebuild the mafia's Provincial Commission. High-ranking actors in the Sicilian Cosa Nostra were generally less well-connected than mid-ranking members and well-connected actors tended to form ties with poorly-connected actors rather than with other well-connected actors and vice versa (Musotto, 2022, p. 82). Further, older mafiosi had fewer connections, were more likely to be in leadership positions, and were more likely to form bridging ties outside of their district to other groups. Studies of the Calabrian `Ndrangheta have similarly found that bosses tend to be well-connected actors that are positioned as brokers, bridging both between otherwise-disconnected mafiosi and between their families and other groups in the organization (Calderoni, 2014, 2016; Calderoni et al., 2017; Calderoni & Superchi, 2019). Prior research suggests that mafias tend to be split into factions, typically around formal organizational divisions into families, that are interconnected by a small number of key bosses who have positioned themselves as brokers.

3.2.3 Analytical Expectations

This research contributes to the literature on the *vory* and on higher-level coordination in mafias in three main ways. First, much of the qualitative research on the structure of the *vory* has focused on the period between the 1970s and early 2000s. Our analysis provides a contemporary view of how the *vory* are structured. Second, prior ethnographic and qualitative accounts of the *vory* have presented a mixed view of their structure, with numerous reported forms and contributing factors. Our analytical approach – outlined in the next section – allows us to explore the structure of the *vory* and disentangle competing explanations. Finally, while prior studies of mafias have discussed the role of bosses, these investigations have largely looked at bosses within the wider organizational structure of mafia families, in large part due to the difficulty of collecting data on the higher-level coordination bodies and mechanisms of mafias. Our unique dataset provides insight into how bosses interact with one another specifically within a high-level association.

3. Investigating the Structure of the Thieves-in-Law

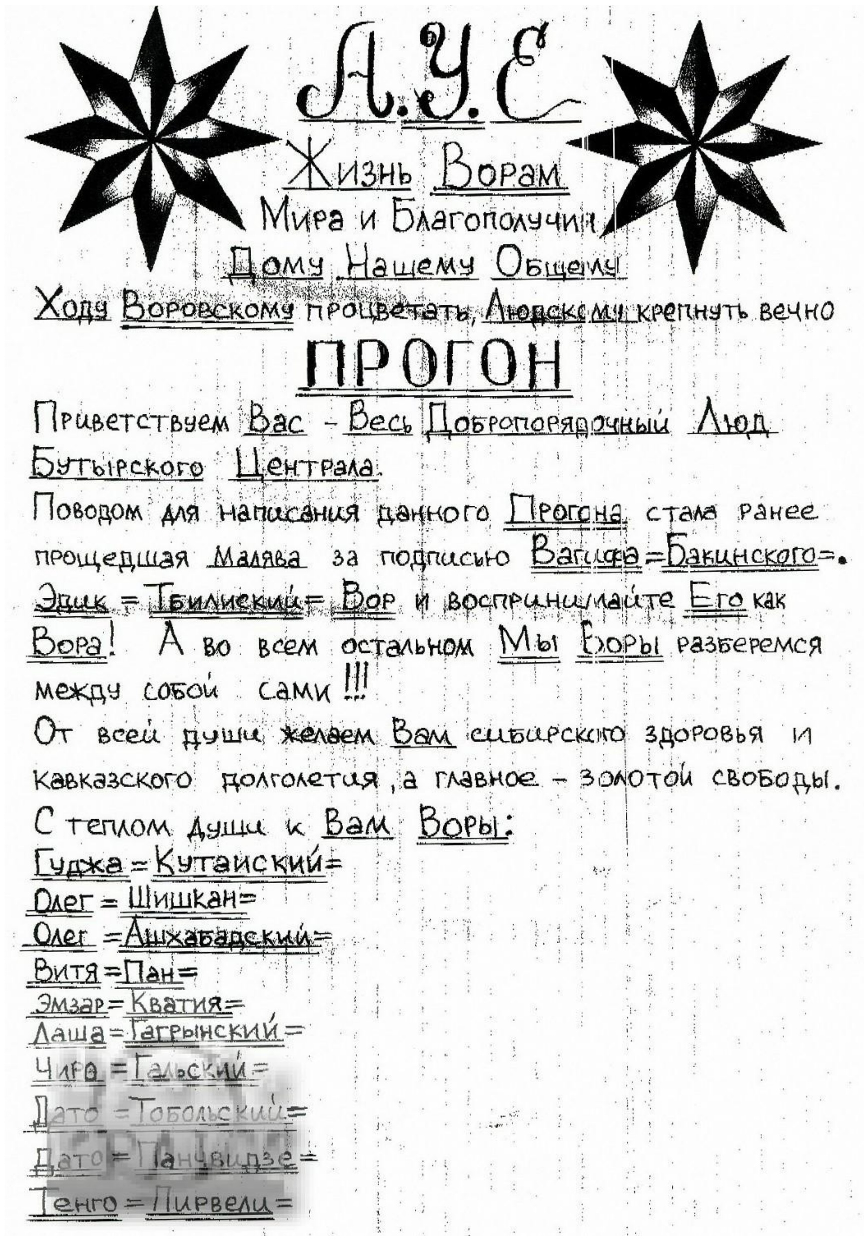
Our analysis explores the structure of the *vory* and is guided by the findings and gaps in the existing literature. Based on accounts of factionalization within the *vory* (Mironova, 2023; Slade, 2014) and within other mafias (Calderoni et al., 2017; DellaPosta, 2017, 2023), we expect that the *vory* will exhibit a clear community structure with densely-connected clusters. Prior research has also suggested that *vory* possess different levels of power and authority (Cowley et al., 2015; Handelman, 1995). We expect that the *vory* will be polycentric, with multiple centers of power forming through the power struggles and competition amongst the different bosses. Based on reports of competition and alliances amongst the *vory* and Musotto (2022)'s finding of degree disassortativity within the Sicilian Cosa Nostra, we test for polycentric rich polarization to see whether (1) the multiple poles of power tend to connect with each other and (2) whether the relationships tend to be especially strong when these key poles do connect. We expect that the factions around these key players will be connected by strategically-positioned brokers similar to the role of certain bosses within other mafias like the Italian-American Mafia and the `Ndrangheta. Finally, prior research on the *vory* has provided two alternative explanations for the emergence of factions: ethnic and generational clustering based on evidence from the 1990s and 2000s (Mironova, 2023; Slade, 2014) and vertical patronage relationships (Galeotti, 2004; Slade, 2014). These competing drivers are difficult to disentangle based purely on qualitative observation. Further, they are not necessarily mutually exclusive: *vory* could seek out patrons or clients of similar age and ethnicity because these similarities could lead to greater trust or could provide more opportunities for *vory* to meet. We test these competing hypotheses to determine which factors are most at play when *vory* form relationships.

3.3 Data and Methods

3.3.1 Data

To investigate these expectations, the present research utilizes a novel dataset of *progony* collected as part of CRIMGOV research grant (PI Federico Varese). As noted in Section 3.2.1, *progony* are a key mechanism for coordination within the fraternity. Further, signing a *progon* is a social decision, impacted by a *vor*'s relationships with the *vory* who have drafted the document and other *vory* mentioned or impacted by its contents, as well as their support for the message itself. As such, investigating the signing decisions of the *vory* can provide insights into their intra-association coordination and their wider social structure. These documents are handwritten, often adorned with elaborate drawings, and use a formulaic language. Figure 3.1 presents an example *progon*.

We have collected a database of 84 *progony* that were published by the Russian news agency PrimeCrime.ru, an online news source focused on the *vory* that has been used in prior studies of the fraternity (Tsereteli, 2022; Varese, 2017; Varese et al., 2021). We filtered the database of 84 *progony* issued from 1988 onwards to focus on those published between 2013-2022; this ten-year period was chosen to ensure that the mafiosi who signed the documents were contemporaneous with one another. We then filtered out texts that were not signed; while edicts traditionally must be signed by those who drafted and supported them, a 2019 law making individuals who sign *progony* liable for prosecution has caused some criminals to publish their documents anonymously (see discussion in Varese, 2024, pp. 48–65). These filters left a sample of 50 signed documents written and published between 2013-2022. We extracted the names of the *vory* who signed the *progony* in the dataset (N = 172).



А.У.Е

Жизнь Ворам

Мира и Благополучия

Дому Нашему Общему

Ходу Воровскому процветать, Людскому крепнуть вечно

ПРОГОН

Приветствуем Вас - Весь Добропорядочный Люд Бутырского Централа.

Поводом для написания данного Прогона стала ранее прощедшая Малява за подписью Вагифа=Бакинского=. Эдик = Тбилиский= Вор и воспринимайте Его как Вора! А во всем остальном Мы Воры разберемся между собой сами!!!

От всей души желаем Вам сибирского здоровья и кавказского долголетия, а главное - золотой свободы.

С теплом души к Вам Воры:

Гуджа = Кутаиский=

Олег = Шишкан=

Олег = Ашхабадский=

Витя = Пан=

Эмзар = Кватия=

Лаша = Гагрынский=

Чиро = Гальский=

Дато = Тобольский=

Дато = Панчвидзе=

Тенго = Пирвели=

Figure 3.1. A photograph of a *progon* distributed in Buturskaya Prison in Moscow in 2014 and signed by 10 *vory* in the bottom-left. Available at https://www.primecrime.ru/photo/2915/.

Using these data, we first constructed a bipartite or 2-mode *vor*-to-*progon* network where an edge from actor *i* to document *j* represents that *i* signed *j*. A bipartite network is made up of two sets of nodes – in this case, the *vory* (N = 172) and the *progony* (N = 50) – that can only form ties with nodes of the other set; in this network, there are no *vor*-to-*vor* or *progon*-to-*progon* edges. We then projected the 2-mode network to create an undirected 1-mode actor network where nodes represent *vory* and edges represent instances where two actors co-signed an edict together; a weighted version based on a count overlap projection – where the weight on an edge represents the number of *progony* that a given pair of *vory* co-signed together – and an unweighted binary version were used for different analyses as appropriate. Summaries of the 2-mode and 1-mode co-signing networks are presented in Table 3.1, while Figure 3.2 depicts visualizations of the networks.

To supplement the main network dataset, we collected additional information from a range of sources. First, we coded the text of the documents for a thematic content analysis of the topics of the *progony*. We reviewed the texts and developed codes for the sections of the documents, then reviewed these codes to generate larger themes. Second, an interview with a former inmate at a penal colony was conducted in December 2022 to provide context on the *vory*, their place within the prison and criminal ecosystem, and the role of *progony*. Finally, we collected additional data on the *vory* within our sample from PrimeCrime.ru. To investigate the role of ethnicity and age in the social structure of the *vory*, we collected the birth year and ethnic background of the actors in our network. Mafiosi were classified as either Georgian (52%), Russian (28%), South Caucasian (13%), North Caucasian (5%), or Eastern European

Table 3.1 *Progony* Co-Signing Network Information

	2-Mode Network	1-Mode Network
Edge Types	*Vor*-to-*progon*	*Vor*-to-*vor*
Total Nodes	222	172
Vory	172	172
Progony	50	*n.a.*
Total Edges	315	2,377
Density	0.04	0.16
Avg. Degree		
Vory	1.83	27.64
Progony	6.30	*n.a.*

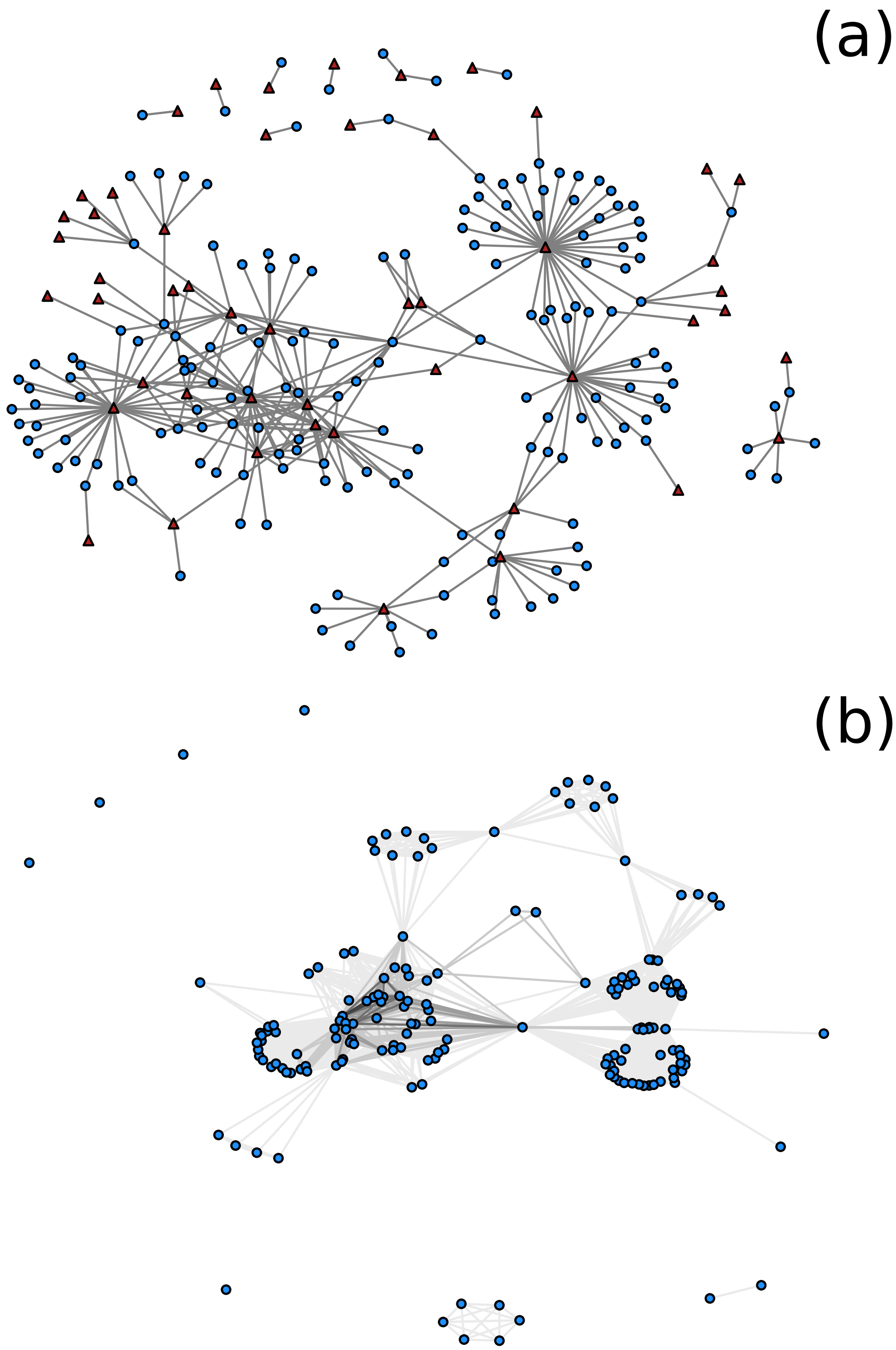


Figure 3.2. Visualizations of (a) the 2-mode *vor*-to-*progon* network where *vory* are represented by blue circles and *progony* are represented by red squares and (b) the 1-mode *vor*-to-*vor* network where darker edges represent higher ties weight.

(2%).[15] Further, information from PrimeCrime.ru was used to construct a matrix of crowning relationships. For each *vor*, PrimeCrime.ru reports which *vory* crowned him and any *vory* that he crowned. The crowning matrix only contains the 172 *vory* involved in signing *progony* in our database; as such, if a *vor* was crowned by another member who had not signed any *progony*, that crowner would not be in the crowing matrix. This matrix contains 667 directed ties from crowners to crownees. Of the 172 *vory* in the network, 97 acted as crowners for other *vory* in the network, while 98 were crowned by at least one other *vor* in the network; overall, 153 actors were involved in crowning ties. In the following analyses, this was operationalized as an undirected and unweighted 1-mode network where an edge represents that one *vor* crowned the other.

3.3.2 Methods

We used a range of analytical techniques to explore the structure of the *vory* fraternity. First, we use descriptive statistics to understand the overall level of clustering in the co-signing network and to detect brokers. We use two measures of bipartite clustering defined by Opsahl (2013). 2-mode reinforcement measures the extent to which two co-signing actors are likely to seek out additional connections together. 2-mode transitivity measures how often actors seek to form direct co-signing ties with the contacts of their connections. These measures range from 0 to 1 and represent the share of potential reinforcement configurations or triadic triples that occur within the observed network. Similar to other studies on mafias, we use betweenness centrality to detect brokers (Calderoni, 2014, 2016; Calderoni & Superchi, 2019; Varese,

[15] 'South Caucasian' includes *vory* from Armenia and Azerbaijan, while 'North Caucasian' includes Chechen, Ingush, and Avar *vory*. 'Eastern European' includes two Moldovan *vory* and one Ukrainian *vor*.

2013b). Betweenness centrality measures the extent to which a given node lies on the shortest path between other pairs of nodes within the network, where high betweenness centrality suggests an actor bridges between different parts of the network.

To explore the presence of factions in the co-signing network in-line with reports in the literature, we used a multiple core-periphery detection algorithm described in Kojaku & Masuda (2017, 2018). This approach extends the original concept of core-periphery structure to allow for multiple core-periphery pairs, where there can be multiple sets of cores that are each densely intra-connected but not well-connected to other cores; each core has an associated periphery of nodes that are connected to their core but not to each other, as well as being poorly-connected to other core-periphery pairs. The number of core-periphery pairs is optimized algorithmically and not set by the researcher. The algorithm functions similar to the Louvain method of community detection, maximizing a modularity function that considers connections within cores, between cores and their peripheries, and between different core-periphery pairs by iteratively reassigning nodes to new core-periphery pairs and between the core and periphery of a given pair. The statistical significance of the multiple core-periphery configuration is determined by comparing the modularity function applied to each core-periphery assignment in the observed network to the values of the function for the same core-periphery assignment of nodes in 500 simulated networks where edges have been randomized using the configuration model. The resulting p-values represent the share of the simulated networks for which the objective function value for a given core-periphery pair is higher than in the observed network. We classify nodes that are within the same core-periphery pair as being part of the same faction. Kojaku & Masuda (2018) find that this algorithm gives highly similar modularity results compared to the Louvain method – a method used to discover communities and factions within other mafias (Calderoni et al., 2017; Catino et al., 2022) – when considering each core-periphery pair as a community. Despite these similarities, this algorithm was chosen over more traditional community detection approaches like the Louvain method to better account for the multiple core-periphery structure suggested by Slade (2014)'s finding that vory clustered around other charismatic or powerful vory.

Prior research on the *vory* suggests the association is informally unequal, with disparities in power and connectivity between different members. Some social networks contain a 'rich club', where prominent or central nodes are well-connected to each other and form a dominant centralized core. To investigate the extent to which central *vory* tend to form ties with each other, we present a range of rich-club coefficient (RCC) distributions. RCC distributions are based upon a given measure of prominence applied to all nodes and measure how well nodes of at least a given prominence level connect to each other; these values are normalized against random networks generated using an appropriate null-model. Normalized RCC value $\phi > 1$ indicates rich-club ordering and $\phi < 1$ indicates that prominent nodes are less connected than expected at random, with a floor of $\phi = 0$. These distributions have been extended to 2-mode and weighted 1-mode networks (Opsahl et al., 2008). The weighted network extension measures "the fraction of weights shared by the rich nodes compared with the total amount they could share if they were connected through the strongest links of the network" (Opsahl et al., 2008, p. 1), while the 2-mode extension calculates the prominence of the nodes in one mode within the bipartite network and then projects to a weighted 1-mode network before calculating the RCCs. Our analysis presents four rich-club distributions: (1) bipartite, with prominence based on the number of *progony* signed; (2) unweighted 1-mode, with prominence based on degree; (2) weighted 1-mode accounting for the strength of ties, with prominence based on degree; and (4) weighted 1-mode, with prominence based on node strength equal to the sum of a node's tie weights.

Finally, we use two hypothesis-testing techniques to disentangle the effects of shared ethnicity, age similarity, and vertical patronage relationships on whether *vory* sign *progony* together. First, we specify a multilevel exponential random graph model (ERGM). ERGMs are multivariate models where the dependent variable is the presence or absence of a tie between two nodes and the independent variables are parameters that represent micro-level configurations in the network like two-stars, triangles, and shared attribute homophily (Hunter, Handcock, et al., 2008; Lusher et al., 2012; Robins et al., 2007). Multilevel ERGMs are an extension of these models to include parameters relating to dependencies both within a single level of a multilevel network and between the levels (Wang et al., 2013). Prior studies on the structure of mafias and other organized crime groups have used both one-level (e.g., Breuer & Varese, 2023; Diviák et al., 2021) and multilevel ERGMs (e.g., Breuer & Baradel, 2025; Coutinho et al., 2020; C. M. Smith & Papachristos, 2016). For this analysis, we combine the 2-mode *vor*-to-*progon* network, which represents co-signing decisions, and the 1-mode crowning network, which represents vertical patronage relationships in-line with prior ethnographic findings. Since our interest is in co-signing and two actors cannot sign a *progon* together until both have been crowned as *vory*, we fix the 1-mode crowning layer – in the simulations underlying the multilevel ERGM, the crowning ties are not allowed to vary. We include parameters relating to the likelihood that two *vory* will sign the same *progon* if they have the same ethnicity, have similar ages, and share a crowning tie; we also include structural parameters relating to centralization and clustering as controls. The multilevel ERGM can be interpreted similarly to a logistic regression because both levels are unweighted; however, interpreting coefficient sizes is difficult, especially when using alternating structural parameters as we do to improve convergence (for technical discussion of alternating structural parameters, see Snijders et al., 2006). Thus, we also use a multiple regression quadratic assignment procedure (MRQAP) model on the weighted 1-mode co-signing network. MRQAP models allow researchers to assess the extent to which a network is influenced by other networks or covariance matrices (Dekker et al., 2007; Krackhardt, 1988); this model is interpreted like a linear regression model where the dependent variable is the strength of the tie between two nodes. These models have been applied to a range of mafias, organized crime groups, and street gangs (Bichler & Norris, 2024; Campana & Varese, 2013; Krajewski et al., 2022). In our MRQAP model, the dependent network is the weighted 1-mode co-signing network; the independent networks are a binary and undirected 1-mode matrix measuring whether two actors share the same ethnicity, a weighted and undirected 1-mode matrix reporting the age difference in years between pairs of actors, and a binary and undirected 1-mode matrix representing crowning relationships. Table 3.2 presents summary statistics for each of the matrices used in the MRQAP model; note that the density of the age difference matrix is 1 because this difference is calculated for all pairs of actors in the network.

These analyses were conducted using a range of software and packages. The 2-mode reinforcement and transitivity, 2-mode RCC distribution, and weighted RCC distributions were calculated using the ‘tnet’ package (version 3.0.16) in R. Betweenness centrality and the 1-mode unweighted RCC distribution were calculated using the ‘NetworkX’ package (version 3.1) in Python, while the Kojaku & Masuda algorithm was implemented using the ‘cpnet’ package (version 0.0.21) based upon NetworkX. UCINet (version 6.758) was used for the MRQAP model, while MPNet was used for the multilevel ERGM.

Table 3.2 MRQAP Matrix Information

	Shared Ethnicity	Age Difference	Crowning Relationships
Weight	Unweighted	Weighted (min. 0, max. 46)	Unweighted
Nodes	172	172	172
Edges	5,362	14,706	667
Density	0.36	1.00	0.05

3.3.3 Limitations and Mitigations

While the data used for the following analyses do face some potential limitations, we have mitigated these as much as possible. PrimeCrime.ru is a trusted and reliable source of information on the *vory*, an otherwise difficult-to-study group. The website has been described as "perhaps the most exhaustive information resource about *vory-v-zakone* available to the public" (Schreck, 2009). Beyond academic researchers and journalists, even criminals themselves use it as a news source: Varese (2017) notes the case of a Russian criminal – whose phone was wiretapped by the police – checking PrimeCrime.ru for information about the recent death of a prominent *vor* (Varese, 2017, p. 93; see also Varese et al., 2021, p. 147). The database only contains documents and signatures whose provenance and veracity have been confirmed on PrimeCrime.ru – *progony* that *vory* have denoted as being fake have been excluded.

There is also the potential limitation of missing data. It is likely that we have not collected all of the relevant *progony* written during the observation period given that *progony* are documents shared amongst member of the post-Soviet prison and criminal community and as such are difficult to access. This is mitigated somewhat by an interview we conducted with a former prisoner who suggested that only approximately 20% of *progony* are related to the fraternity, while 80% are focused on local day-to-day life for common prisoners (C1); we are only interested in the edicts that relate to the structure of the *vory*, which is the set that PrimeCrime.ru collects. Despite this, however, there are likely still missing relevant documents. This is a limitation that cannot be easily remedied by triangulation from other available data sources or interviews. The potential missing data limits the generalizability of our results to general coordination amongst the *vory* and to other mafias. However, *progony* signing decisions are often driven by underlying social relationships between *vory* and relate to the

Table 3.3 *Progony* Published per Year

Year	Progony Published
Unknown	6
1955-2012	21
2013	14
2014	7
2015	1
2016	4
2017	4
2018	5
2019	4
2020	4
2021	11
2022	3
Total	84

coordination of decisions within the association. These data, while not perfect, are still suitable for achieving our aim of exploring the social structure of this higher-level coordination body. Further, we lack reliable geographic information for the *vory*, which could have an impact on co-signing decisions; this impact could be direct, where mafiosi in the same location have greater opportunity to physically sign together, and indirect, where criminals who have spent time together either in or out of prison may have formed personal or professional ties that could lead to co-signing. While PrimeCrime.ru does provide some geographic information about *vory*, it lacks many observations; as such, it was excluded from the analysis. The exclusion of geographic information is mitigated by the fact that mafiosi do not need to be together to co-sign a document, as edicts can be physically transported or sent by text message between criminals during the signing process before being distributed (Varese, 2017, 2024).

3.4 Results

3.4.1 Progony Sample Descriptive Statistics

The following tables provide descriptive statistics about the dataset of *progony* used in the network analysis. Table 3.3 shows the number of *progony* issued by year within the database. Note that Table 3.3 also includes 21 *progony* that were issued prior to 2013, 6 *progony* for which a publication year could not be determined, 6 *progony* published after 2013 that were not signed, and 1 *progon* that was written by a non-*vor*; these *progony* have been excluded from the network analysis. 2013 and 2021 were the two most active years for *progony* publication, representing 25% and 19% of progony written between 2013-2022.

Table 3.4 presents the number of *vory* by how many *progony* they have signed, while Table 3.5 presents the number of *progony* by how many signatures they received. Note that

Table 3.4 Count of *Vory* by Number of *Progony* Signed

Number of *Progony* Signed	Count of V*ory*	%
1	104	60%
2	35	20%
3	12	7%
4	8	5%
5	8	5%
6	2	1%
7	3	2%
Total	172	100%

Table 3.5 Count of *Progony* by Number of *Vory* Signing

Number of *Vory* Signing	Count of *Progony*	%
1	24	48%
2	6	12%
3	0	0%
4	3	6%
5-9	6	12%
10-19	6	12%
20+	5	10%
Total	50	100%

these tables only consider *vory* and *progony* used in the network analysis. 60% of the 172 *vory* in our network only signed one *progon*, while a further 20% signed two *progony*. Similarly, nearly half of the *progony* were only signed by one actor; on the other hand, approximately 35% of *progony* had more than 5 signatures, with 22% having 10 or more. These results suggest there are centralizing tendencies in the network – a small number of *vory* sign many *progony* and *progony* with many signatures appear to attract more signatures than those with only a few signers. One potential explanation for centralization in signatures is preferential attachment, where *progony* with many signatures attract further signatures. This could arise from *vory* wanting to sign documents that have already been signed by many other *vory* to show their support and appear relevant; further, powerful *vory* are likely able to marshal the support of many of their clients, allowing them to publish well-signed *progony* that appear to have support from many of the fraternity's members.

Table 3.6 summarizes the topics covered in the 50 *progony* used in the network analysis. Note that a given document is able to cover multiple topics. Cancelling or 'de-crowning' a *vor* is the most common topic, coming up in 54% of *progony*. De-crowning plays two important roles amongst the *vory*. First, it allows the association to police its borders, removing members who do not contribute to the community, who act in a way inconsistent with the thieves' code, or who were potentially fraudulently admitted. Second, it is a key way in which powerful *vory* compete and vie for power – cancelling the supporters of a rival removes their allies from a position of power. Other popular topics include general discussions of the *vory*'s laws and status (10%) and changes to the rules (8%).

Table 3.6 Count of *Progony* by Topic

Topic	Count of *Progony*	%
Cancellation of the status of a *vor*	27	54%
General speculations about *vory*'s laws and contemporary situation in the fraternity plus invectives against specific members	5	10%
Changes of laws or setting new rules	4	8%
Support for a specific *vor* or specific *vor*'s emissaries	4	8%
Greetings to a new *vor* arriving to a prison	4	8%
Appointments of new *vory*'s emissaries	3	6%
Requirement of actions (punishment of rule violators or cancellation of crownings)	3	6%
Announcement that a membership of a *vor* is 'under conflict'	3	6%
Information about new crownings	3	6%
General discussion against the use of the Internet and social media	2	4%
Problems with or prohibition of gambling in prisons	2	4%
Prohibition of drugs in prisons	1	2%
Total Progony	50	

3.4.2 Exploratory Network Analysis

To better understand the social structure of the *vory*, we first focus on the extent to which there is clustering and factionalization within the co-signing network. We find preliminary evidence of moderate clustering within the 2-mode *vor*-to-*progon* network using Opsahl (2013)'s clustering measures. The reinforcement and transitivity scores for the 2-mode network are 0.23 and 0.26 respectively. *Vory* tend to both reinforce their co-signing decisions by signing other documents with the same co-signers and close triads in approximately 25% of cases where it is possible to do so. While exhibiting a trend towards clustering, *vory* also create ties outside of their clusters.

To better understand the boundaries and structures of these clusters, we apply the multiple core-periphery detection algorithm described in Kojaku & Masuda (2018). Figure 3.3 presents a visualization of the largest connected component of the unweighted 1-mode co-signing network where nodes are colored by their core-periphery pair, taken as topological factions; the largest connected component contains 92% of the nodes (N = 158) and 99% of the edges (E = 2,361).

Each of the factions shown in Figure 3.3 is statistically significant at the $p < 0.05$ level based on a statistical test using randomly-generated networks created using the configuration model as the null-model; eight of the nine detected groupings are significant at the $p < 0.001$ level. This suggests that the observed faction structure is stronger than what would be expected at random. The findings on clustering are observable in the visualization: the members of a given factor are densely connected to each other, but not to members of other factions. Through their *progony* co-signing decisions, the *vory* segment themselves into densely-intraconnected clusters that are generally disconnected from other topological factions in the network.

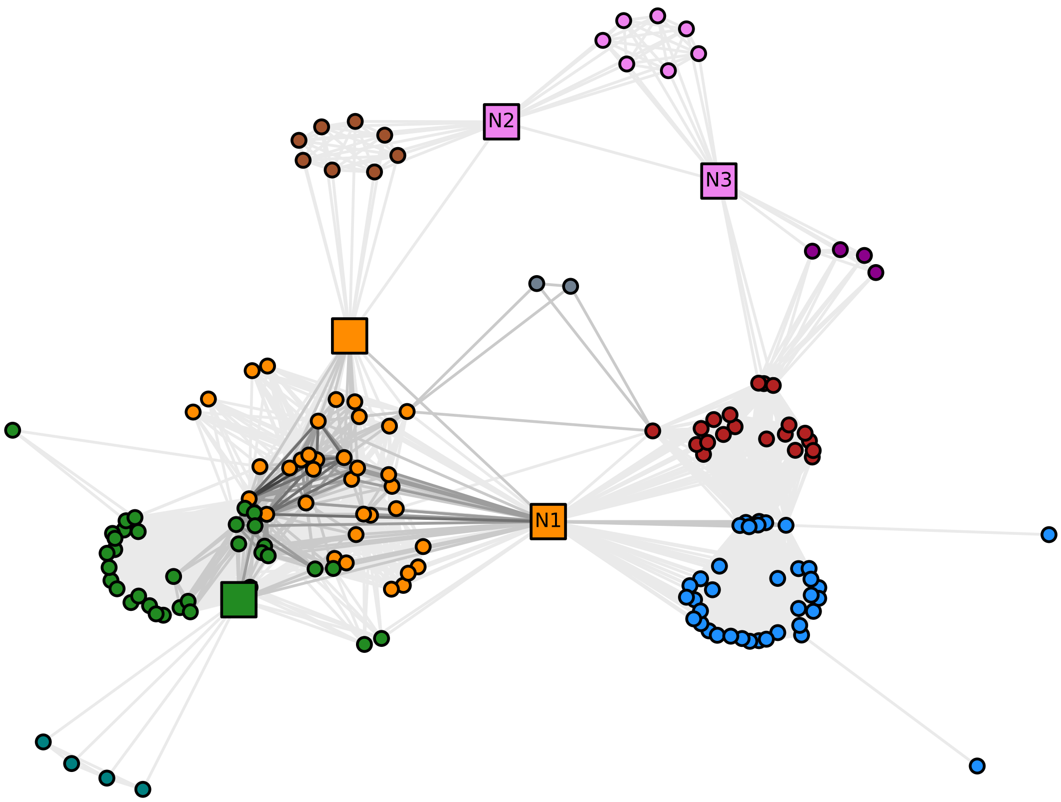


Figure 3.3. Visualization of the largest connected component of the 1-mode vor-to-vor network. Node colors represent faction affiliations, while the large square nodes represent the five nodes with the highest betweenness centrality scores.

There are, however, two key ways in which factions are connected. First, some factions share close relationships with each other, as evinced by strong interconnections between multiple members. For example, among the four largest factions, there are two faction pairs where each faction shares material overlap with its counterpart, suggesting close social and working relationships that would enable joint competition against other factions. Second, individual *vory* also act as key bridges between different factions. In Figure 3.3, the five actors with the highest betweenness centrality scores – a measure of node centrality often used to identify brokers in social networks – are denoted as squares. The central node labeled N1 acts as a particularly important broker among the *vory*, with a betweenness centrality scores three times higher than the next-highest actor. This *vor*, named Niko Dgebuadze, is not a prominent or powerful actor. Instead, he is a member of multiple factions, as evinced by his direct co-signing relationships with members from four factions, and mediates exchange between the different groups. He was arrested twice for drug possession but spent as little as a year in prison. Hence, he was able to coordinate the *vory*'s communication from outside of prison and signed seven *progony* between 2013-2022. According to PrimeCrime.ru, he was a protégé of another more prominent *vor* – Gela Kardava – and "established himself as a selfless fighter" for his protector's interests during numerous conflicts among clans (PrimeCrime.ru, 2013). The shortest paths between the four largest factions go directly through Dgebuadze. The other brokers identified as square nodes in Figure 3.3 play a similar role, although to a lesser extent than Dgebaudze. For example, *vory* Oleg Shishkanov (N2) and Konstantin Borisov (N3) signed two *progony* each, one of them together. The two were close friends and partners and crowned other *vory* together (Vershov, 2018). Both were involved in numerous alliances within the fraternity and thus shorten paths between members of different factions. This second form

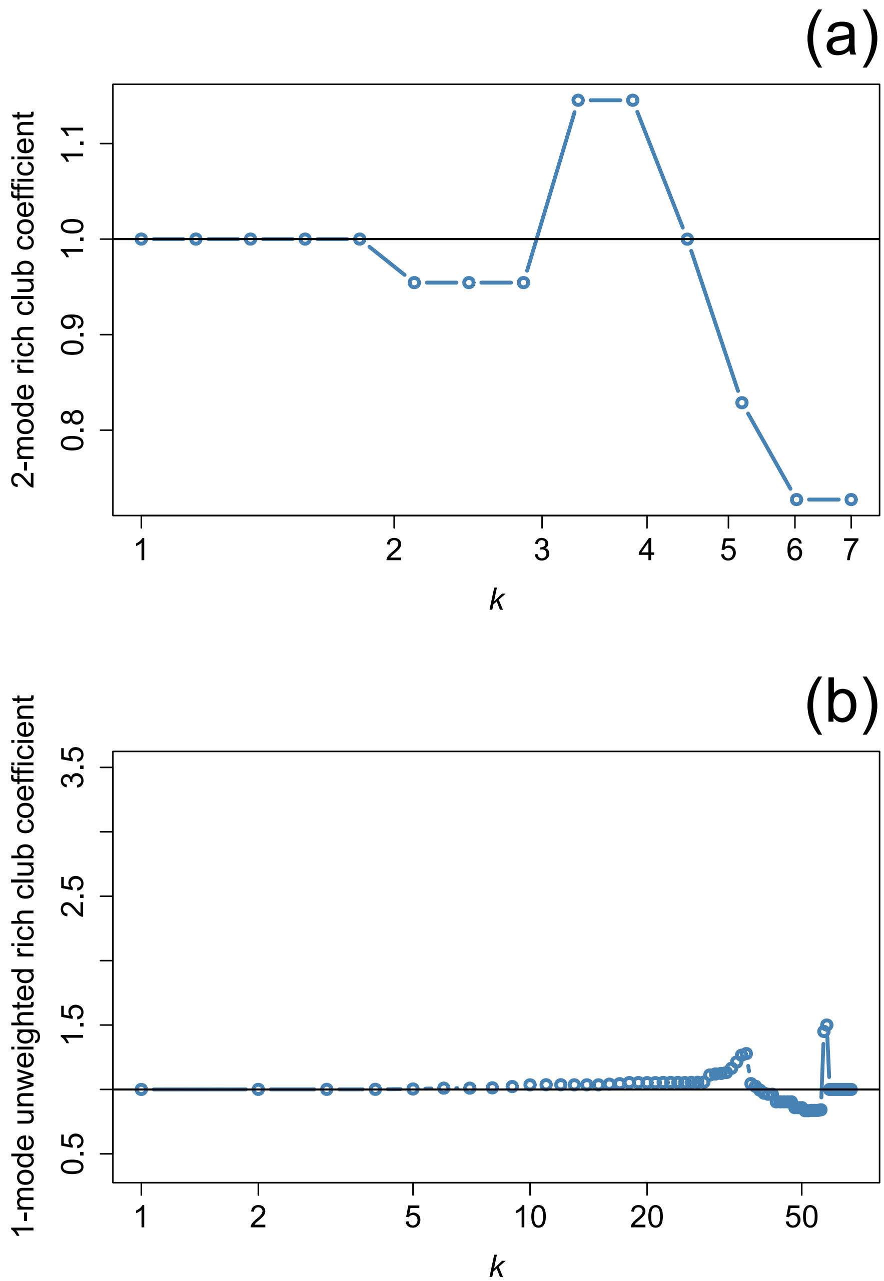


Figure 3.4. Rich-club coefficient distributions for (a) the 2-mode *vor*-to-*vor* network where prominence is determined by the number of *progony* a *vor* has signed and (b) the unweighted 1-mode *vor*-to-*vor* network where prominence is determined by a *vor*'s degree *k*. Horizontal axes are presented on a logged scale.

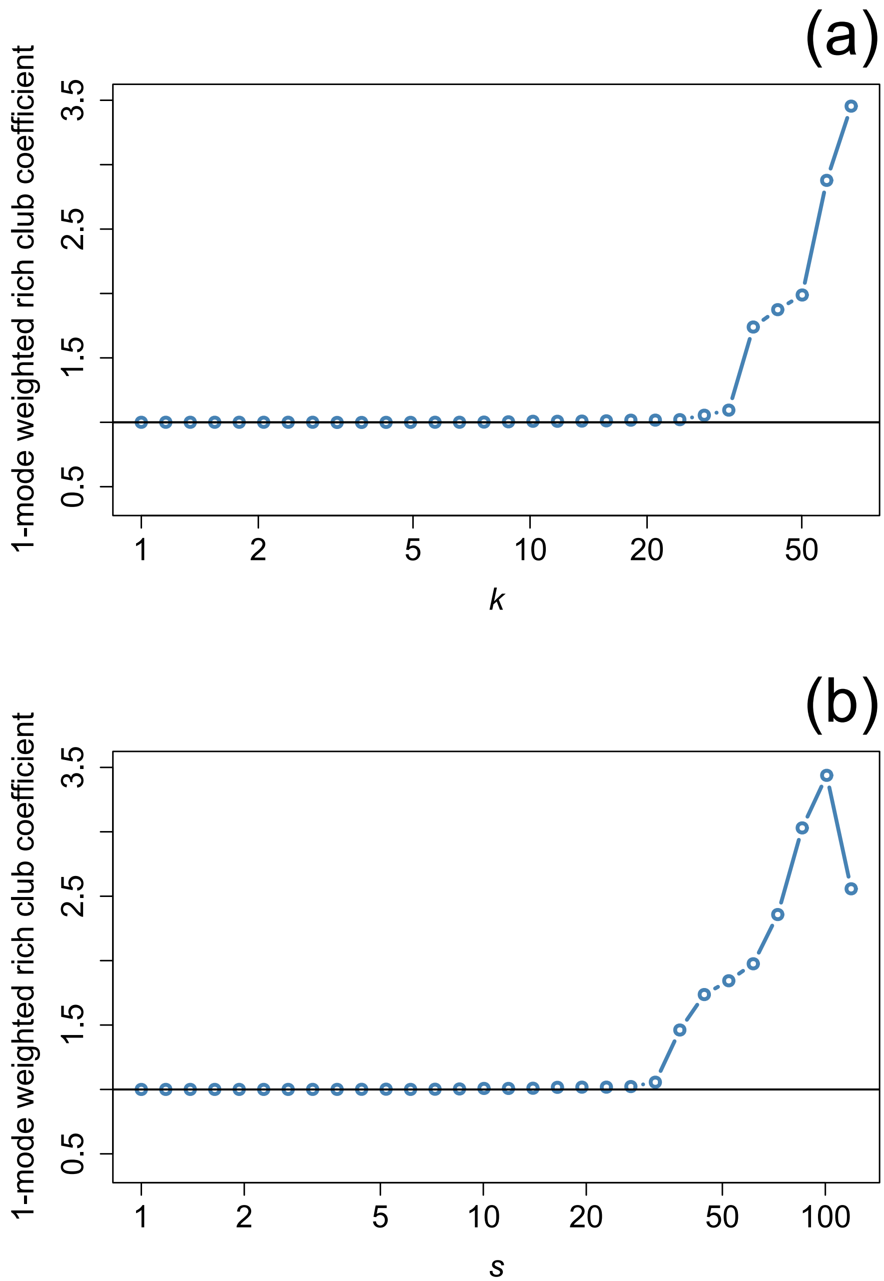


Figure 3.5. Rich-club coefficient distributions for the weighted 1-mode *vor*-to-*vor* network where prominence is determined by (a) a *vor*'s degree *k* and (b) a *vor*'s strength *s*. Horizontal axes are presented on a logged scale.

of interconnection enables cohesion on the part of the entire association; while less indicative of close working relationships than overlap, these brokers enable communication and cooperation between otherwise-disconnected clusters.

Results from the rich-club coefficient distributions suggest the *vory* association shows signs of polycentric rich polarization. Figure 3.4 shows the RCC distributions for the 2-mode and unweighted 1-mode networks. The negative effect shown in the 2-mode RCC distribution suggests that *vory* that sign more *progony* tend to co-sign together less than expected at random; the effect size, however, still indicates some level of connectivity between frequently-signing *vory*. Conversely, the unweighted 1-mode RCC distribution suggests the co-signing network does not show strong topological rich-club ordering – nodes with many co-signers connect approximately as often as expected at random.

While the unweighted RCC distribution suggests that well-connected *vory* do not have a material tendency to connect with each other, the distributions in Figure 3.5 suggest that when prominent *vory* do connect, they tend to form stronger ties than expected at random. This is true when prominence is defined as an actor's number of co-signers or the strength of the ties with those co-signers. *Vory* that sign many documents and those that co-sign with many others are not more likely to co-sign together, leading in multiple centers of power rather than one club of 'rich' nodes; when these 'rich' nodes do connect, however, they form strong ties.

3.4.3 Hypothesis-Testing Model Results

Beyond exploring the social structure of the *vory*, we also test competing hypotheses from the literature on the drivers of connections between *vory*. First, results from the multilevel ERGM are presented in Table 3.7. The structural controls confirm some of the findings from

Table 3.7 Multilevel Exponential Random Graph Model Results

Parameters	Coefficient	SE	t-ratio	SACF
Edge Statistic	-11.788*	0.658	0.057	0.310
Structural Effects				
Progony Popularity	4.339*	0.367	0.049	0.247
Vory Co-Signing Reinforcement	-0.032	0.018	-0.023	0.032
Attribute Effects				
Age Difference	-0.001	0.001	-0.006	0.044
Shared Ethnicity	0.080*	0.015	-0.006	0.081
Crowning Relationship	0.540*	0.050	0.028	0.117

Statistical significance indicated by *. Parameters for which the absolute value of the ratio between its coefficient and standard error is greater than 2 are considered significant.

Sections 3.4.1 and 3.4.2. The *Progony* Popularity parameter is positive and significant, suggesting that *progony* with many signatures are more likely to attract additional signers. The term for 2-mode reinforcement is negative, small, and statistically insignificant. The attribute parameters suggest that *vory* of the same ethnicity and who share a crowning tie are more likely to co-sign together. Heuristically, the Crowning Relationship coefficient is larger than the value for the Shared Ethnicity term, suggesting that crowning relationships are likely a more impactful driver of co-signing than shared ethnicity. However, as noted in Section 3.3.2, interpreting and comparing the magnitudes of ERGM parameter coefficients can be difficult; to determine the relative importance of ethnicity, age, and crowning relationships in co-signing decisions, we specify a series of MRQAP models.

Models 1 and 2 in Table 3.8 exclude shared ethnicity and crowning relationships respectively, while Model 3 contains all three relevant parameters. All three models do not find a significant effect of age similarity on co-signing behavior. While the coefficients for shared ethnicity in Models 1 and 3 are positive and statistically significant, the effect size is small; further, the R^2 value for Model 1 – excluding the crowning relationships parameter – is much lower than the values for Models 2 and 3 which do include crowning relationships. Some effect of ethnic homophily is to be expected given prior findings in the literature on the network structure of organized crime groups (Oliver et al., 2014). The small effect size and low R^2, however, suggest that ethnicity-based clustering is not a material determinant of the observed network structure. The crowning relationship parameter, on the other hand, is both statistically significant and a much stronger determinant of co-signing relationships; its coefficient size is nearly nine times larger than that for shared ethnicity and the parameter's inclusion in the model drives most of the explained variance based on the increase in R^2. Further, the coefficient

Table 3.8 MRQAP Results

	Model 1		**Model 2**		**Model 3**	
Variables	Coefficient	SE	Coefficient	SE	Coefficient	SE
Intercept	0.178***	0.000	0.182***	0.000	0.161***	0.000
Shared Ethnicity	0.060***	0.017	--	--	0.058***	0.017
Age Difference	-0.001	0.001	-0.002	0.001	-0.002	0.001
Crowning Relationship	--	--	0.513***	0.026	0.512***	0.026
Model Fit						
R^2	0.004***		0.052***		0.056***	
R^2 (adj.)	0.004***		0.052***		0.055***	
***p<0.01, **p<0.05, *p<0.1						

values for the shared ethnicity and crowning relationship parameters are largely similar across models regardless of whether the other parameter is included; this suggests that the parameters are not collinear and that they are independent. Rather than *vory* crowning new recruits of the same ethnic group, the crowning relationships appear to cut across the ethnicity cleavage, with *vory* showing no material observed preference for those from their ethnic group.

For the analyses presented in this section that utilized the weighted 1-mode network – the RCC distributions and MRQAP model – the results are robust to the projection method used. Appendix C includes results for these two analyses using the Newman projection for 2-mode networks (Newman, 2001) which mirror the findings in Section 3.4 using the count overlap projection.

Table 3.9 presents the goodness-of-fit results for the multilevel ERGM. For terms included within the model, t-ratios with absolute values below 0.1 suggest the model has converged well; for global network metrics, t-ratios with absolute values below 2.0 suggest the model fits that feature well (Wang et al., 2009, 2013). The t-ratios for both within-model terms and global network metrics meet these criteria, suggesting the model fits the observed network well. These results suggest the model was well-behaved, non-degenerate, and tended to capture the structural characteristics of the observed network well.

3.5 Discussion

The findings presented in Section 3.4 depict the *vory* as a multiethnic, multigenerational association structured around factions but interconnected by a small number of key brokers. Overall, members cluster around multiple poles of power that are poorly connected to each other, creating polycentric polarization; when these poles connect, however, they tend to form strong ties. These findings of polycentricity and factionalization are in-line with prior research.

Table 3.9 Multilevel Exponential Random Graph Model Goodness-of-Fit Results

Parameters	Observed	Mean	St. Dev.	t-ratio
Within-Model Terms				
XEdge	315.000	315.897	18.863	-0.048
XASB	485.007	486.371	34.786	-0.039
XACB	2540.938	2546.905	282.657	-0.021
Age X2StarBDifference	20486.000	20511.392	3442.688	-0.007
Ethnicity X2StarAMatch	1208.000	1213.636	212.495	-0.027
ATXAX	408.250	406.002	31.232	0.072
Global Network Metrics				
stddev_degreeA[1]	7.923	7.923	0.000	-1.000
skew_degreeA[1]	1.735	1.735	0.000	-1.000
clusteringA[1]	0.187	0.187	0.000	-1.000
stddev_degreeX_A	0.968	0.991	0.093	-0.257
skew_degreeX_A	1.109	1.229	0.254	-0.473
stddev_degreeX_B	8.180	8.470	0.662	-0.439
skew_degreeX_B	2.304	2.360	0.431	-0.131
clusteringX	0.066	0.073	0.013	-0.527

[1] These global network metrics apply to the 1-mode crowning network level that was fixed during estimation and thus did not vary.

Despite this, the fraternity is more cohesive than prior qualitative accounts suggest. Different topological factions share interconnections among multiple members. This fragmented system is further held together by a small number of *vory* that have relationships across multiple factions. Similar to the department-bridging managers in Burt (2004)'s study or Italian-American Mafia bosses that connect disparate families in DellaPosta (2017, 2023) – both of which bridge formal divisions – a small set of *vory* act as brokers between the informal factions identified in the co-signing network. While the *vory* are factionalized, they maintain interconnectivity that can allow for cooperation, coordination, and reconciliation.

Prior accounts disagree on the basis for factionalization. We find that factions are not ethnically- or generationally-defined; instead, they are more associated with crowning relationships and vertical patronage ties that cross-cut ethnic and generational cleavages. This fact has likely contributed to the historical and present resiliency of the *vory* fraternity. Unlike other mafias, the *vory* face unique challenges because they are not monoethnic and contain bosses with a wider range of ages than other mafias. These factors present cleavages along which the association could split in times of internal conflict or external pressure, leaving the fraternity vulnerable to schisms. Cross-cutting institutional relationships like crowning and patronage, however, act to buttress against the tendency for actors to form ties with similar others that can result in disconnected homogenous communities. While it is unlikely that the *vory* intentionally designed this structure, it likely played a part in maintaining the fraternity through decades of state repression and violent conflict. Despite this, viewing the *progony* co-signing network reveals the *vory* system to be fragile and vulnerable to fragmentation generally. While vertical patronage relationships can strengthen against demographic attribute-based factionalization, the network is still based on strongly intraconnected communities tenuously held together by a few key brokers.

These findings can provide insight into the high-level structures of other mafias. While the ability to generalize our results is limited by the specific nature of the data and differences in context and organizational structure between the *vory* and other mafias, trends in recent years and decades suggest that other mafias may have become more similar to the *vory*. The higher-level coordinating bodies of many mafias like the Sicilian Cosa Nostra, the Italian-American Mafia, and the Hong Kong and Taiwanese triads have become defunct due to increased police pressure; even for those where formal coordination mechanisms still exist – like the `Ndrangheta and the yakuza – increased law enforcement attention in the past two decades has weakened their efficacy. Further, this police pressure has made regular meetings attended by all bosses or representatives more difficult. The lack of formal coordinating bodies and growing logistical problems are similar to the problems facing the *vory*. As they become more disparate and lose formal coordination structures, our findings suggest that mafias may face factionalization as weaker bosses seek the protection of more influential ones; this further creates opportunities for brokers to bridge structural holes and connect not just between families but also between clusters of bosses.

3.6 Conclusion

Despite internal conflict and external pressures, the *vory* have existed for more than a century. Prior qualitative evidence on the structure of the fraternity has been mixed and contradictory. Our exploratory analysis based on their *progony* co-signing decisions finds that the *vory* today are fragmented but interconnected. Key poles of power draw other members towards them using vertical patronage relationships. Clusters are largely disconnected, but a small set of strategically-positioned brokers bridge between groups. Finally, we find that co-signing decisions are primarily driven by patronage relationships – operationalized using crowning ties – while ethnicity and age are not material determinants of co-signing. The *vory* are a largely multiethnic, multigenerational, polycentric association that is both factionalized and interconnected.

Our findings contribute to the literature on the *vory* by providing insight into the underlying social structure of the association. Further, we contribute to the literature on mafia structure by providing a first look inside a mafia's higher-level coordination body. Despite limitations around potential missing data, a lack of geographic data, and differences in context, our results provide insight into how bosses in other mafias may interact, especially as these other mafias increasingly face the same problems as the *vory* like the inability to meet and the destruction of formal higher-level coordination bodies. Finally, our findings have implications for the formation of factions and organizational structure in non-criminal organizations; while the *vory* face particular issues as criminals, legal-market organizations like corporations and legislative bodies also face challenges from informal power hierarchies, attribute differences, and the growth of siloed subgroups.

Future research can build upon our findings in three main ways. First, our analysis focuses on *progony* co-signing, which is one way in which *vory* interact; future studies can look to collect additional relational information on the *vory* to better understand the whole social structure of the fraternity. Second, our research suggests that using non-standard data sources can enable analysis of the structure of higher-level coordination within other mafias, presenting a future avenue for better understanding other mafias and comparing between them. Finally, our findings on factionalization and brokers are similar to those of Burt (2004)'s analysis of mid-level managers in a large electronics company; future research should look to compare cluster formation and brokerage across different types of organizations. For example, the role of brokers may differ between legal and illegal settings and between settings where factions are organizationally-defined or arise from informal social relationships. The story of the *vory* provides insights into how factions can arise informally and how organizations can remain resilient and enable coordination.

4

Testing the Reliability of OSINT Network Data for Investigating Organized Crime Infiltration of Legal-Market Businesses

Abstract[16]

This paper assesses the usability of open-source intelligence (OSINT) – in particular, public business registers – for investigating organized crime group infiltration of legal markets. While OSINT can be easier to access than data from traditional closed sources, there are doubts about its reliability. To investigate the utility of this new data source, data were collected from an Italian pre-trial detention order and the Italian business register to construct networks representing an `Ndrangheta group involved in penetrating legal-market businesses. Results suggest that OSINT offers several advantages over traditional closed sources and works as a proxy when considering local network structures. However, OSINT network data face challenges when operating as a proxy for closed-source police data networks to identify key

[16] A version of this paper has appeared in: Breuer, N. (2025). Testing the reliability of OSINT network data for investigating organised crime infiltration of legal-market businesses. *Global Crime*, 1-25. https://doi.org/10.1080/17440572.2025.2567277

actors and reproduce global metrics. OSINT network data from business registers are valuable for investigating infiltration of legal-market businesses, although researchers must use them to answer questions to which they are well-suited.

4.1 Introduction

Beyond their primary criminal activities like drug trafficking and racketeering (Calderoni, 2012; Sciarrone & Storti, 2014), organized crime groups often seek to generate additional earnings, gain power, and launder their criminal proceeds by infiltrating legal businesses (Berlusconi, 2016; Riccardi, 2014; Savona, 2015). To do so, they create new businesses, invest in existing companies, or partner with legal entrepreneurs, either voluntarily or involuntarily. Europol reports that 80% of organized crime groups in Europe use legal business structures in the course of their criminal activities, with nearly half infiltrating existing businesses or founding their own (Europol, 2021). Academic research has begun to investigate this phenomenon in the past decade. Prior studies have largely focused on the types of companies infiltrated, the impact of organized crime involvement in legal businesses, and new methods for identifying suspicious companies. The structure of criminal groups as they penetrate and exploit legal businesses, on the other hand, has been understudied. The focus has largely been on the companies rather than the criminal actors. A better understanding of this understudied aspect of organized crime is crucial for investigating and disrupting these groups. This gap in our understanding is due, in part, to a lack of reliable and rich data on the ties between criminal actors as they look to infiltrate legal businesses.

One potential data source is based on open-source intelligence (OSINT) from public business registers, official lists of companies within a country that report information about companies' financial results and ownership. Ownership information can enable academic researchers and law enforcement to create corporate interlock networks that represent which actors sit on the boards of the same companies together. These business registers offer a low-cost and easy-to-access alternative to the closed-access data sources typically used to investigate organized crime group structure like wiretap transcripts and indictments. Despite these benefits, however, the nature of OSINT means that it may be less reliable than data collected from more commonly-used sources. This is especially true when investigating organized crime groups, whose members often intentionally hide their activities and relationships from public view. This research sought to validate the reliability of OSINT network data for investigating the structure of organized crime groups as they infiltrate legal businesses. To do so, network data from public business registers were compared to closed-source police data from an indictment, the sort of data previously used throughout the literature and largely viewed as reliable. This research aimed to answer two primary research questions:

1. Are network data from publicly-available business registers useful for investigating the structure of organized crime groups, either alongside or apart from closed-source data?
2. Can network data from publicly-available business registers serve as a reliable proxy for harder-to-access data from more commonly-used closed sources?

To answer these questions, closed-source network data were collected from an Italian pre-trial detention order. The case related to an anti-mafia operation focused on a Calabria-based `Ndrangheta group. The group was active in both Italy and Germany and was involved in criminal activities including extortion, money laundering, and anti-competitive practices like forcing businesses out of commercial areas through threats of violence. The Italian business register, Il Registro Imprese, was then used to collect corporate interlock network data for the companies within the indictment.

Comparing the OSINT and closed-source data reveals that business register network data hold numerous benefits over using only closed-source police data, including greater detail and full inclusion of all relevant actors. Further, a number of instances were identified where the police data appear to be inaccurate based on the official documents. Business register network data are valuable both as a standalone dataset and as a necessary supplement to closed-source police data where they are available. This is especially important for contexts where official closed-source data may be non-existent or difficult to access. Considering OSINT sources' usability as a proxy for police network data, the business register network data did not reliably produce similar node-level or global results to police data. However, they did produce similar local neighborhood-levels results, suggesting that network data from business registers can be used as a proxy for analyses that focus on local network dynamics. These findings have implications for academic researchers, law enforcement, and policymakers. Professionals in these fields would benefit from greater use of the lower-cost but still reliable data available in public business registers, but must also be wary of the applications for which these data appear not to be well-suited. This research contributes to the existing literature by expanding the sorts of data sources available for analysis and validating the use of business register data when investigating organized crime groups, with a particular focus on which sorts of research questions and analyses are best served by using network data from business registers. Using this novel data source opens up the study of new sociological and criminological questions about how organized crime groups operate and structure themselves in legal markets.

The article is structured as follows. Section 4.2 reviews the relevant literature on infiltration of legal businesses and the use of OSINT for investigating organized crime groups. Section 4.3 introduces the data and methods. Section 4.4 presents the results of the data source comparison and network analyses. Section 4.5 discusses the results and Section 4.6 concludes by discussing the implications of the findings.

4.2 Literature Review and Background

4.2.1 Literature Review

Research into the structure of organized crime groups has largely focused on the relationships between criminal actors as they carry out their primary criminal activities, ranging from racketeering and the provision of private protection to drug trafficking and human trafficking. These studies generally draw their data from one of three sources. First, many studies investigating the structure of organized crime groups in Italy use pre-trial detention orders, judicial files presented to a judge to request permission to arrest suspects in advance of a trial (Calderoni, 2012, 2014; Calderoni et al., 2017; Calderoni & Superchi, 2019; Musotto, 2022; Tumminello et al., 2021; Varese, 2013b). These files are typically rich in details and evidence to allow the judge to justify their decision to allow for pre-trial detention and contain information on the suspects as well as the alleged criminal activities. Further, they often contain direct evidence from wiretaps and physical surveillance conducted during the investigation and information on the structure of the organized crime group. Second, other studies make use of other judicial sources like prosecutorial files (Bright et al., 2019; Catino et al., 2022; Natarajan, 2000, 2006) and penal records (Tumminello et al., 2021). Finally, some studies use data directly from law enforcement agencies and policing systems (Baika & Campana, 2020; Coutinho et al., 2020; Diviák et al., 2021; Morselli & Roy, 2008). All three sources are non-public and difficult to access, typically requiring a relationship with prosecutors, law enforcement, or other members of the criminal justice system. These sources are then used to collect network data on direct relationships between criminal actors like communication (e.g., Bright et al., 2019; Calderoni, 2012, 2014), collaboration and co-offending (e.g., Baika &

Campana, 2020; Coutinho et al., 2020), and intermarriage and kinship ties (e.g., Catino et al., 2022; Tumminello et al., 2021).

Beyond the core criminal activities of organized crime groups, investigations of infiltration of legal businesses have largely focused on the companies being infiltrated. Studies have investigated the industries and geographies most likely to be infiltrated (Caneppele et al., 2009; Gurciullo, 2014; Kruisbergen et al., 2015; Mirenda et al., 2019; Riccardi, 2014; Riccardi & Maggioni, 2025; Savona, 2015; Sciarrone & Storti, 2014). These studies suggest that infiltration is most likely to occur in companies in traditional organized crime group strongholds and sectors like wholesale and retail trade, construction, and hospitality. Other studies have focused on the impact of organized crime infiltration, finding evidence of diversion of public funds by organized crime groups (Barone & Narciso, 2015; Checchi & Polo, 2020; Gennaioli & Tavoni, 2016) and negative effects on other nearby firms (Calamunci et al., 2022; Calamunci & Drago, 2020; Mirenda et al., 2019). Finally, another group of studies has focused on developing risk assessment approaches for identifying suspicious companies and countries (Aziani et al., 2022; Garcia-Bernardo et al., 2017; Jofre, 2022; Jofre et al., 2024). Some of these approaches use open-source ownership data from corporate intelligence databases like Orbis to construct networks connecting firms to their beneficial owners and then look for structural characteristics that can be a sign of infiltration or financial crime (e.g., Jofre, 2022; Jofre et al., 2024). While helpful for identifying potentially-infiltrated companies, these ownership networks often take the form of egonets, with each firm connected to its ownership chain of shareholders and beneficial owners, both individuals and other businesses. The focus is on the firms themselves rather than the interconnected social structure of the actors involved in the companies. This research differs from these prior approaches by considering the whole-network structure of organized crime groups as they attempt to infiltrate legal businesses, rather than the ownership structure within individual businesses.

In contrast to closed-source data where access must be granted to researchers, OSINT is based on the collection and synthesis of data from publicly-accessible sources (Chainey & Alonso Berbotto, 2022). OSINT is further distinguished from open data, which consists of previously-collected datasets that have been shared for further use in research (Open Knowledge Foundation, n.d.). OSINT consists of new data collection. In recent years, investigations of infiltration of legal businesses and of organized crime more broadly have started to use OSINT that can offer both a lower-cost alternative to closed-source data and a new perspective potentially absent from those closed sources. As noted above, a number of studies investigating financial crime and the presence of organized crime groups in legal markets have made use of national business registers (e.g., Mirenda et al., 2019) and corporate intelligence databases like Orbis (e.g., Aziani et al., 2022; Jofre, 2022; Jofre et al., 2024). Other studies on organized crime have collected data from social media sites and other online sources (Carley, 2015; Dolliver et al., 2018), public media reporting (Chainey & Alonso Berbotto, 2022; Krebs, 2002), official government data (Chandra et al., 2014; Checchi & Polo, 2020; Gurciullo, 2014), publicly-available satellite images (Pinto Hidalgo & Silva Centeno, 2022), and biographies of criminals (Morselli, 2001, 2003). OSINT data, however, can be less reliable than data from closed sources (Carley, 2015; Chainey & Alonso Berbotto, 2022), requiring additional evaluation and validation.

To investigate the structure of organized crime infiltration of legal businesses, this research draws on approaches developed for investigating the connections between corporate boards. Corporate interlock refers to cases where a member of one company's board of directors is also a member of another company's board. Studies of corporate interlock typically use publicly-available and open-source information on the board members of corporations collected from business registers, stock markets, or corporate intelligence databases (S. Battiston & Catanzaro, 2004; Benischke et al., 2024; Davis et al., 2003; Glattfelder & Battiston, 2009; Kogut & Walker, 2001). Interlocking directorates can enable inter-company coordination and embed companies within a wider social structure (Davis, 1996). While previous research on infiltration of legal businesses has used similar sources and considered ownership information, these studies have analyzed the ownership structure within siloed companies. Corporate interlock analyses, on the other hand, focus on how the overlap between ownership structures can act as a channel for information diffusion or represent underlying social structures.

4.2.2 Data Source Background

Before laying out the specifics of the empirical case used in the following analyses, this section first focuses on the relative strengths and weaknesses of the available sources. Commonly-used closed sources like pre-trial detention orders, judicial files, and police data have two primary strengths over OSINT sources. First, they typically offer richer, more detailed, and potentially more reliable data on organized crime groups than OSINT sources can (Carley, 2015; Diviák & Lord, 2024). These sources are the result of months or years of investigation and work by law enforcement, prosecutors, and other members of the criminal justice system. They often draw from a wide base of evidence including wiretaps, physical surveillance, arrest records, and interviews with both suspects and victims (Campana & Varese, 2022).

Second, these sources can provide information on aspects that cannot be seen through purely open sources because they are collected using a rigorous process aimed at uncovering criminal activities. Based on the illegality of their activities, criminal actors have an incentive to hide and obscure their actions, especially from the highly-visible sources from which OSINT draws (Carley, 2015). This is especially true of the data used to study infiltration of legal businesses, which are collected from business registers and corporate intelligence databases. These sources draw from government systems based on self-reporting and registration by companies. They are likely to see manipulations by criminals who report distorted financial results or false ownership information (Riccardi & Berlusconi, 2016). Criminals may hide their ownership of a company behind 'clean' figureheads or proxies who show up as the official owner instead of the criminal actor. While this relationship can be discovered through surveillance and police investigation, it would not be visible based solely on network data from publicly-available business registers – the figurehead would appear in the register while the criminal actor would not. Further, business registers alone do not contain information on which actors are part of an organized crime group, partnering directly with an organized crime group, or unknowledgeable about the infiltration at all.

On the other hand, business registers have a number of benefits over these commonly-used closed sources, both as a supplement to closed-source data and as a standalone data source. An overarching challenge with using closed-source data from the policing and judicial systems is that these sources are not constructed with academic research in mind. While rich in detail, these documents and files are curated by law enforcement and prosecutors for the purpose of investigating and ultimately convicting the members of organized crime groups. This aim leads to priorities that may differ from those of academic researchers and biases around what data are collected and ultimately included in the documents (Bright et al., 2021; Campana & Varese, 2022). For example, wiretapped conversations between organized crime group members are purposively sampled for inclusion in law enforcement data sources, with police and prosecutors only including conversations and actors they deem to be relevant for their investigation or case (Berlusconi, 2013). The inclusion of the entire universe of affiliated actors is not always necessary for the case being made, but is important for academic research involving network analysis: while some network centrality measures have been shown to be robust to missing data from purposive sampling in some cases (Berlusconi, 2013), the observed characteristics of a network are generally sensitive to the inclusion or exclusion of specific nodes or edges (Campana, 2016). While business registers are also not curated with academic research of this kind in mind, these OSINT sources allow researchers to gather all of the relevant data based on search terms and a structured data collection approach and then make boundary specification decisions around the inclusion and exclusion of certain nodes that fit their research design (Carley, 2015).

Network data from publicly-available business registers also have the benefit of being relatively low cost and easy to collect for both law enforcement agencies and academic researchers. Business registers and corporate intelligence databases are pre-existing sources that are designed to be easily searchable (Jofre et al., 2024). For law enforcement, business registers can provide a data source that does not require them to go through the time-intensive and expensive process of electronic and physical surveillance. This aspect is especially valuable in the early stages of investigations for allocating scarce resources by, for example, prioritizing surveillance subjects. For academic researchers, business register data can provide an alternative to negotiating access to closed-source data. Further, business registers can enable research in contexts and countries where closed-source official data is either inaccessible or non-existent.

The accessibility and inexpensiveness of business registers, while beneficial, is only important if the resulting data are useful for research. First, the business register data can act as a supplement when a researcher has access to an indictment or other closed source (Campana & Varese, 2022). OSINT can function as a triangulation tool to give greater confidence in the closed-source data and clarify uncertainties (Carley, 2015). Second, business register data can also be valuable as a standalone dataset depending on the research question and research design. While data from business registers are unable to identify covert company affiliation ties, the official ties they do identify are sociologically and criminologically relevant for understanding the interlocking corporate system created by organized crime groups as they infiltrate legal businesses. As discussed in more depth in Section 4.3.1, corporate interlock ties are indicative of which actors are involved in the infiltration and subsequent management of a legitimate business. The actors who form official or covert ties with a company participate in that company's operations. These ties are especially significant since illegal-market actors are willingly making themselves and their ties to other companies and actors visible in official government records. Further, while closed sources like indictments can contain evidence from a wide range of sources, business register documents contain comprehensive and rich detail about infiltration specifically.

These data also include information on network dynamics. Business register documents contain information on which actors are affiliated to which companies, as well as when the affiliation began and ended (InfoCamere, n.d.). Police and judicial sources, on the other hand, do not necessarily contain complete temporal information. Prior research has found that ignoring the timing and sequence of ties can obscure network dynamics and underlying social structures (Falzon et al., 2018). In the case of corporate interlock networks, the timing, duration, and sequence of affiliations can impact conclusions. For example, a static tie representing that two actors were both affiliated to the same business obscures whether the actors joined the company at the same time or whether one actor 'followed' the other. Similarly, static ties cannot distinguish actors who founded a company from those who joined at a later date. Further, aggregating company affiliations over a long timespan into static ties can make it difficult to distinguish whether two actors were associated with the same company at the same time or whether one actor left before another joined. This temporal view of company ownership and involvement also extends beyond the timeframe of law enforcement data, which typically only include information collected during a police investigation from a 'closed' case prior to disruption (Diviák & Lord, 2024).

Finally, business registers also contain non-network information about companies including geospatial data, industry, corporate structure, financial results, and company-to-company information on mergers and acquisitions (InfoCamere, n.d.). While commonly-used closed sources may mention some of these data points, they are only included when deemed immediately relevant to the case by prosecutors, similar to the purposive sampling of actors and wiretapped conversations. These features mean that data from business registers are particularly well-suited for answering research questions that focus on the entire system of criminally-infiltrated companies and all related actors rather than just the network of mafiosi. For example, a combination of affiliation-based network data and company attribute data can be used to investigate industry specialization within mafia networks in legitimate sectors – are actors industry-agnostic or do they specialize by only participating in companies within a given economic area? Do specialists in the same industry cluster together in sector-specific communities and, if so, what role do brokers play within a fragmented system? Further, business register data can enable studies that contribute to our understanding of the determinants of organized crime group structures. Prior research suggests that the fact that an organized crime group like a mafia engages in illegal activities (Haller, 1990; Reuter, 1983) and the specifics of the chosen criminal activity (Breuer & Varese, 2023) can impact that group's structure. One way of contributing to this debate would be to compare the structure of a mafia group as its members infiltrate legitimate businesses – based on business register data – against the same group's structure as it commits illegal activities like racketeering or drug trafficking, as seen in closed-source data. When investigating these questions, business registers can offer richer and more granular data than closed sources.

4.3 Data and Methods

4.3.1 Data

To investigate the reliability of network data from publicly-available business registers, data were collected from both open and closed sources. Closed-source data came from an Italian pre-trial detention order relating to Operation Stige. Operation Stige targeted a Calabrian `Ndrangheta clan and its associates and resulted in the arrests of 169 actors in Italy and Germany in January 2018. Beyond presenting evidence of extortion, attempted murder, and corruption, the pre-trial detention order for the operation contained detailed accounts of the mafiosi's attempts to both found and infiltrate legal businesses across a range of economic sectors. These accounts were based on evidence from wiretaps and physical surveillance of meetings between suspects.

This document was coded to extract a list of companies that had been infiltrated by the `Ndrangheta clan. This list included all companies mentioned in the indictment to which an

`Ndrangheta member or associate was officially or covertly affiliated. Official company affiliations refer to associations that would show up in a company's ownership information such as founder, owner, shareholder, or executive. Covert ties refer to unofficial, hidden ownership of or involvement in a company. For example, a mafioso may have invested in a company by providing a 'clean' and apparently-legitimate accomplice with capital to invest in the business. In this case, the accomplice would be officially affiliated to the company, while the mafioso would be covertly associated with the business. The mafioso holds a de facto share of the company but not in such a way that would show up in the company's ownership structure.

The company names in this initial list were searched in the Italian business register, Il Registro Imprese published by the Camera di Commercio. Official documents were obtained for each company in the initial list that could be identified in the register. These documents contained a range of information on Italian companies, including the identities of the officers and shareholders over time, their tax ID, and their address (InfoCamere, n.d.). This information, alongside matching data in the pre-trial detention order, was used to disambiguate and deduplicate the initial company list. Comparison against the business register documents revealed that 11 of the company names on the initial list were aliases or were misreported in the indictment, leading to duplicates. After removing these duplicates, 165 total companies were identified in the indictment. Of these, 33 could not be identified in the business register to download their documents. 20 were based in Germany and not enough information was provided to identify them in the German business register, while one company was mentioned but unnamed. For the remaining 12, the pre-trial detention order named the companies and noted they were based in Italy but did not provide enough identifying information to find their documents in the business register. Further, in the course of consulting the register, three companies owned by relevant actors were discovered that were not reported in the indictment and were subsequently included in the final company name list.

These data were used to create two corporate interlock datasets. The first, the indictment-based corporate interlock dataset, contained the 165 companies identified from the pre-trial detention order and the 242 actors reported as being affiliated to those companies. The dataset consisted of relational ties between actors and companies. Each tie was tagged as official or covert depending on the relationship the actor had to the business. Qualitatively, these corporate interlock ties between actors and companies were the main operational ties of interest when looking at mafia infiltration of legal markets. Based on the evidence provided within the pre-trial detention order, actors with an official or covert corporate interlock tie to a given company were the ones who managed the day-to-day operations and illegal activities of that company. Further, actors who were affiliated to the same companies worked together. Corporate interlock indicated cooperation in the management of businesses in the legal market and enabled inter-company coordination. For example, a number of forestry companies were connected by shared corporate interlock ties that enabled collusion when bidding for public contracts.

The second, the OSINT-based corporate interlock dataset, contained both the 132 businesses from the pre-trial detention order for which business register documents could be located and the 3 additional relevant companies identified during data collection. This dataset also included the 291 actors mentioned in the business register documents as being officially affiliated to the businesses as officers or shareholders. No covert ties were included, as these were not visible in the underlying data source. All actors mentioned in the documents were included, not just those that were also mentioned in the pre-trial detention order. While the business register information included temporal data on when actors joined and left companies, this data was excluded from the dataset for comparability with the indictment-based corporate interlock dataset. While the pre-trial detention order did include some information on when company affiliations occurred, this data was not available for the majority of affiliations and was mentioned inconsistently.

A third dataset was collected by coding the indictment for instances where actors communicated with each other either over the phone or in person based on wiretap records and physical surveillance reports. This dataset contained actor-to-actor ties. Each communication tie was further tagged as either being business-related, where the topic of conversation related to the infiltration of legal businesses, or non-business-related. To avoid distortions from the inclusion of extraneous and irrelevant people, actors were removed from the indictment-based communications dataset if they satisfied three criteria: (1) they were not present in the indictment-based corporate interlock dataset, (2) they were not charged with a crime in the pre-trial detention order, and (3) they were not involved in any business-related communications.

These three datasets were then used to create eight main networks, summarized in Table 4.1. The indictment-based corporate interlock and OSINT-based corporate interlock datasets were used to create corporate interlock networks. These datasets were used directly to construct 2-mode networks in which an edge represents an actor's affiliation to a company. To investigate how actors collaborate and connect, these 2-mode networks were projected to create 1-mode actor-to-actor networks. In these networks, edges between actors represented that two actors were affiliated to the same company. To better understand the structural differences between the main components of these networks, core versions were also created that removed components of fewer than four nodes that could potentially impact results. Finally, the

Table 4.1 `Ndrangheta Corporate Interlock and Communications Networks Information

	OCI 2-Mode	*ECI 2-Mode*	*OCI Full 1-Mode*	*OCI Core 1-Mode*	*ECI Full 1-Mode*	*ECI Core 1-Mode*	*Comms. (All Ties)*	*Comms. (Business Ties)*
Network Type	2-Mode	2-Mode	1-Mode	1-Mode	1-Mode	1-Mode	1-Mode	1-Mode
Edge Type	Actor-to-Company	Actor-to-Company	Actor-to-Actor	Actor-to-Actor	Actor-to-Actor	Actor-to-Actor	Actor-to-Actor	Actor-to-Actor
Total Nodes	426	407	291	209	242	158	212	189
Actors	291	242	291	209	242	158	212	189
Companies	135	165	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Total Edges	377	405	795	756	482	454	599	431
Density	0.01	0.01	0.02	0.04	0.02	0.04	0.03	0.02

indictment-based communications dataset was used to create two communications networks. One network includes only communication ties that related to the infiltration of companies, while the other contains all communication ties.

The networks based on the indictment-based corporate interlock dataset are denoted as 'ECI', while the networks based on the OSINT-based corporate interlock dataset are referred to as 'OCI'. 1-mode corporate interlock networks that contain all components are referred to as 'Full', while the subgraphs that exclude small components are called 'Core'. The 2-mode corporate interlock networks and communications networks did not have small components removed. The communications networks are denoted as 'All Ties' or 'Business Ties' based on the content of the communications used to generate the edges.

Projecting from 2-mode networks to 1-mode networks can introduce biases or reveal otherwise-hidden structural properties based on the method and weighting used to generate 1-mode edges. To account for this, the analyses were conducted based on three different weightings. The binary weighting creates unweighted ties between actors that were affiliated to at least one company together. The count overlap weighting (CO) creates ties between actors with a weight equal to the number of companies to which they are both affiliated. Finally, the Newman weighting projection (NW) is based on Newman (2001) and gives greater weight to co-affiliation at a company that has fewer affiliated actors.

4.3.2 Methods

Beyond the heuristic assessment of the strengths and weaknesses of open- and closed-source data presented in Section 4.2.2, the information available in each source was compared to answer the first research question about the usability of network data from business registers. The overlap between the four full 1-mode networks was calculated. These metrics report what percent of actor nodes were shared between pairs of 1-mode networks, as well as what percent of edges between shared nodes were also present in both networks. These results quantify how much information from one network is captured by another. As noted in Section 4.2.2, OSINT data can also be used to triangulate and confirm the accuracy of closed-source data. To test this, official ties between actors and companies that were in both the 2-mode OCI and 2-mode ECI networks were compared. This analysis identified ties in the 2-mode ECI network that were erroneous based on the official business register reporting.

To address the research question about the reliability of network data from publicly-available business registers as a proxy for closed-source data, the OCI networks were compared against the ECI and communications networks at three levels of analysis: node-level centrality rankings, global metrics like centralization and clustering, and local neighborhood structural similarities. Beyond the utility of business register data in and of itself for answering specific questions, these data may also be useful as a proxy for harder-to-collect closed-source data. Although business registers may not be able to provide direct evidence of hidden activities like covert company affiliation or communications between criminals, the OCI network may share close structural similarity to the ECI and communications networks. If so, this could have value to both law enforcement agencies, who could use OSINT network data to allocate resources and guide surveillance targeting in the early stages of investigations, and academic researchers, who could use network data from business registers to represent underlying criminal networks and investigate their structures and dynamics.

First, node-level centralities were calculated for the actors in each network to determine whether central actors in the OCI network were also central in the ECI and communications networks and vice versa. Nodes in the unweighted networks were compared based on their degree, betweenness centrality, closeness centrality, and eigenvector centrality. Weighted analogues for degree and eigenvector centrality were used for the weighted networks.

The whole-network structures were then compared. The measures of 2-mode clustering defined by Opsahl (2013) were used to compare the 2-mode OCI and ECI networks. 2-mode reinforcement is defined by the number of 4-cycles divided by the number of 3-paths and represents multiple co-affiliation. This metric measures whether two actors affiliated to the same company are likely to be affiliated to multiple other companies together. 2-mode transitivity extends triadic closure to bipartite networks and is defined as the number of closed 4-paths divided by the number of 4-paths, measuring whether actors seek to form co-affiliations with the contacts of other actors to whom they are co-affiliated.

1-mode networks were compared across four metrics. Degree centralization ranges between 0 and 1 and captures the extent to which degree is equally distributed among nodes; a value of 0 means that all nodes in the network have the same degree, while values approaching 1 mean that a small number of nodes are involved in increasing shares of the network's edges. Transitivity and average local clustering coefficient represent the extent to which open triads in the network are closed, forming complete triangles between three nodes. Both metrics also range from 0 to 1. Finally, degree assortativity measures the extent to which nodes are inclined to form ties with other nodes of the same degree. Degree assortativity ranges from -1 to 1. Positive values mean nodes tend to connect with other nodes with similar degrees, while negative values mean high-degree nodes connect to low-degree nodes and vice versa. Weighted analogues of average local clustering coefficient and degree assortativity were used to compare the weighted networks.

These aggregate network-level metrics, however, can obscure underlying differences or similarities in local-level structures. As such, the structural similarity between the networks was also tested using the NetSimile algorithm, a network similarity metric that compares networks by looking at the distribution of local egonet features for each node. This allows for comparison of network structures with greater granularity than that offered by blunt global metrics. Formulated for unweighted networks by Berlingerio et al. (2012, 2013), NetSimile calculates seven features for the egonet of each node in each network: (1) number of nodes in the egonet, (2) local clustering coefficient, (3) average neighbor degree, (4) average neighbor local clustering coefficient, (5) number of edges in the egonet, (6) the number of outgoing edges from the egonet, and (7) the number of neighbors of the egonet. The distribution of each feature across all nodes is then summarized in a vector by its median, mean, standard deviation, skewness, and kurtosis. The Canberra distance between the vectors from the two networks being compared is then calculated and normalized, with a score of 0 meaning the networks are not at all similar while a score of 1 means the networks are identical (Berlingerio et al., 2012, 2013). This research updated the algorithm for application to weighted networks by using weighted analogues for all features aside from the number of nodes in the egonet and the number of neighbors of the egonet.

4.3.3 Limitations

Several limitations based on the data and analyses used in this study should be acknowledged. First, the collection of the open-source business register documents used in the following analyses was based on closed-source information. Relevant companies were first identified in the pre-trial detention order and then searched for in Il Registro Imprese. This means that the data collection process was not fully open-source. However, there are multiple potential ways of obtaining an initial ‘seed’ list of relevant companies. Rather than full access to closed sources that include relational data, partial closed-source information may be more accessible for researchers, such as a list of infiltrated or suspected companies. In contexts in which any closed-source information is unavailable, other methods could be used such as graph mining techniques to identify suspicious companies (Aziani et al., 2022; Jofre, 2022; Jofre et al., 2024) or the use of public media reporting or official government reports (Chainey & Alonso Berbotto, 2022; Krebs, 2002). Future research is required to validate that these methods can provide valid lists of companies on which to collect data from business registers. The focus of the analyses presented in the following section was to determine whether business register documents on companies, once identified, could function as a supplement to or proxy for relational data from closed sources.

Second, the generalizability of the results may be limited. The analysis was based on one case study of a specific organized crime group in a particular institutional context. The `Ndrangheta may infiltrate legal markets differently than other mafias specifically or other organized crime groups more generally. This means that business register-based corporate interlock network data could be more or less reliable for other groups than for the one considered in this paper. Further, business registers in countries other than Italy may be less useful than Il Registro Imprese for the purpose of studying organized crime group structure. Some countries do not have business registers, while others have a large number of local or regional registers rather than one with a national scope. Further, other business registers may not be as complete as the Italian database. They could cover fewer companies or contain less information on those companies, including who the officers and shareholders of a business are and when they joined the firm.

Third, the process of collecting data from both closed sources and business registers can also lead to limitations. The list of companies from the pre-trial detention order used as the basis of both the OCI and ECI networks could be incomplete. Mafia-infiltrated companies may not have been discovered by police or may not have been mentioned in the indictment because they were not deemed relevant for the criminal case brought against the group. This would lead to systematic missingness, where whether or not a piece of data is missing is based on the characteristics of that data. Based on the nature of the underlying source, the boundary and coverage of the network datasets are determined by those of the law enforcement data (Berlusconi, 2013; Campana & Varese, 2022). This limitation affects all of the network datasets. Further, the pre-trial detention order and the business register documents were manually coded for both corporate interlock and communications data. This introduced the potential for subjectivity to bias the data. This limitation was mitigated by collecting information on all companies and conversations mentioned in the pre-trial detention order, reducing the likelihood that particular data would be overlooked or excluded unnecessarily. The communications data were then filtered for relevance using the systematic inclusion rules outlined in Section 4.3.1. This was based in part on a subjective assessment of whether conversations were business-related. Finally, triangulating between multiple data sources posed data collection issues around matching the names of companies and actors between sources. The pre-trial detention order included additional identifying information like the tax ID numbers and addresses of businesses and dates of birth of actors, allowing for disambiguation and matching against the Italian business register information.

Table 4.2 Shared Nodes Between 1-Mode `Ndrangheta Networks

	OCI Full	*ECI Full*	*Comms. (All Ties)*	*Comms. (Business Ties)*
OCI Full	100%	62%	32%	29%
ECI Full	74%	100%	56%	53%
Comms. (All Ties)	43%	64%	100%	89%
Comms. (Business Ties)	45%	68%	100%	100%

4.4 Results

4.4.1 Shared Network Data Between Sources

The following results quantify the difference between the OCI network and the ECI and communications networks. As noted in Section 4.3.1, 33 of the companies in the 2-mode ECI network (20%) were not identifiable in the business register and so not included in the 2-mode OCI network; similarly, three companies in the OCI network (2%) were not included in the pre-trial detention order. Table 4.2 presents the overlap in actor nodes between the 1-mode networks. Values represent the share of nodes from the network in the row that are also present in the network in the column. The OCI network did a better job of capturing the actors from the ECI and communications networks than vice versa, although the overlap between the OCI and ECI networks and the communications networks was generally low.

Table 4.3 presents edge overlap results. Note that in each comparison between networks, only edges that contain nodes present in both networks were considered. For edge overlap, the ECI and communications networks captured more of the edges in the OCI network than vice versa. The full 1-mode OCI network captured approximately 73% of edges from the ECI network for which the nodes are present in both networks. The OCI network, however, covered a materially small share of the communication ties.

4.4.2 Inaccuracies in Closed-Source Data

Further comparison of the OCI and ECI datasets revealed a number of instances in which the pre-trial detention order was incorrect. To identify these cases, covert ties in the ECI dataset were excluded to allow for comparison with the OCI data. Only the 132 companies and 179 actors that were part of both datasets were considered. This was done to account for the fact

Table 4.3 Shared Edges Between 1-Mode `Ndrangheta Networks

	OCI Full	*ECI Full*	*Comms. (All Ties)*	*Comms. (Business Ties)*
OCI Full	100%	91%	39%	41%
ECI Full	73%	100%	49%	51%
Comms. (All Ties)	24%	32%	100%	80%
Comms. (Business Ties)	24%	37%	100%	100%

Note: Only considers edges between nodes that appear in both networks that the indictment-based corporate interlock dataset omitted a number of actors who were affiliated to companies based on their business register documents. The aim of this analysis was to look for discrepancies even in ties between actors and companies common to both data sources. In total, 17 companies (13%) were identified for which the indictment-based data does not match the official business register data. 11 of these cases involved an actor mentioned elsewhere in the pre-trial detention order not being noted as being affiliated to a particular company. For example, actor A was affiliated to company X in the indictment but was not mentioned as affiliated to company Y, while the business register documents for companies X and Y revealed that actor A was in fact affiliated to both companies. Four cases involved the indictment reporting an actor as being officially affiliated to a company and the business register data not including this affiliation, while the remaining two cases involved the two sources reporting the same number of affiliated actors for a company but disagreeing on at least one of the actors. These were cases where the indictment either provided incorrect information or failed to mention an affiliation between a relevant actor and a relevant company, given that both the actor and company were mentioned elsewhere in the document.

4.4.3 Node-Level Centrality Comparisons

To investigate the reliability of network data from the business register for proxying closed-source data, this section focuses on node-level centralities. One way in which network data from business registers could be useful as a proxy is if they were able to identify actors who were central in the ECI and communications networks. To determine whether this is the case, a range of centrality rankings were calculated for each network. The most central actors in the ECI and communications networks were then compared to the rankings for the OCI networks; note that comparisons occurred between networks with the same weighting schemes and the

Table 4.4 Unweighted Central Nodes Comparison Between 1-Mode `Ndrangheta Networks – Top 10

	Degree Centrality		**Betweenness Centrality**		**Closeness Centrality**		**Eigenvector Centrality**	
	In OCI Nodelist	*Overlaps OCI Top 10*	*In OCI Nodelist*	*Overlaps OCI Top 10*	*In OCI Nodelist*	*Overlaps OCI Top 10*	*In OCI Nodelist*	*Overlaps OCI Top 10*
ECI Full	80%	10%	70%	30%	60%	10%	60%	0%
ECI Core	80%	10%	70%	30%	50%	10%	60%	0%
Comms. (All Ties)	30%	0%	10%	10%	30%	0%	20%	0%
Comms. (Business Ties)	20%	0%	20%	10%	10%	0%	20%	0%

Table 4.5 Weighted Central Nodes Comparison Between 1-Mode `Ndrangheta Networks – Top 10

	Weighted Degree Centrality		**Weighted Eigenvector Centrality**	
	In OCI Nodelist	*Overlaps OCI Top 10*	*In OCI Nodelist*	*Overlaps OCI Top 10*
ECI (Full, CO)	80%	10%	100%	0%
ECI (Full, NW)	70%	40%	100%	0%
ECI (Core, CO)	80%	10%	70%	0%
ECI (Core, NW)	70%	40%	60%	0%

ECI Core network was compared to the OCI Core network, while all other comparisons occurred between full networks. The analysis first checked whether the top nodes for each centrality metric for each network were in their corresponding OCI network analogue. It then checked how many of the top nodes for the ECI and communications networks were also at the top of the centrality ranking for the corresponding OCI network. Tables 4.4 and 4.5 present the results for the top 10 for unweighted and weighted centrality measures respectively.

The business register network data tended to include the top ECI nodes, with 50-80% of the top 10 ECI nodes being captured within the unweighted OCI networks and 60-100% being captured within the weighted OCI networks. The top nodes in the communications networks, on the other hand, tended not to be represented in the business register network data. Only 10-30% of the top 10 communications nodes were also part of the OCI networks. Even when the business register network data did contain the top nodes from the ECI and communications networks, these nodes tended not to be central within the OCI networks. The OSINT network data failed to correctly identify important nodes in the networks that contained non-public criminal ties.

Appendix D contains the results for the top 25 nodes by each centrality metric and confirms these results. The network data from the business register did not contain many central nodes in the ECI and communications networks. Even when these important nodes were shared between the networks, they were not central within the OCI networks. As such, business register network data were not a suitable proxy for identifying key actors in closed-source data.

4.4.4 Global Network Metric Comparisons

Similarly, the OCI network struggled to capture whole-network structural characteristics of the ECI and communications networks. Looking first at the bipartite networks, the 2-mode ECI network exhibited materially stronger clustering than the 2-mode OCI network. The 2-mode reinforcement and 2-mode transitivity values for the ECI network were 0.31 and 0.22 respectively, compared to 0.15 and 0.04 for the OCI network.

Global metric results for the unweighted 1-mode networks are presented in Table 4.6. While all of the networks exhibited similar levels of degree centralization, the OCI networks were more clustered and exhibited stronger degree assortativity. Both the Full and Core OCI networks showed high levels of clustering and degree assortativity, while the Full and Core ECI networks showed more moderate clustering and no marked tendency for or against assortativity. The communications networks, on the other hand, had similar global metrics with the lowest levels of clustering and negative values for degree assortativity.

Global metric results for the weighted 1-mode OCI and ECI networks in Table 4.7 show similar structural differences. Note that for the purpose of understanding the differences between the networks, comparisons should only be made between equivalent pairs of OCI and ECI networks. For example, the OCI (Full, CO) and ECI (Full, CO) should be compared to each other and not to other networks. Based on these comparisons, the OCI networks exhibited greater weighted clustering than their ECI analogues. Further, the OCI networks with the count overlap weighting showed moderate degree assortativity, while the Newman weighting networks had degree assortativity values near zero. The ECI networks, on the other hand, generally exhibited small and negative degree assortativity values. Overall, the OCI network

Table 4.6 Unweighted Global Metrics Across 1-Mode `Ndrangheta Networks

	Degree Centralization	*Transitivity*	*ALCC*	*Degree Assortativity*
OCI Full	0.10	0.88	0.74	0.55
ECI Full	0.11	0.62	0.57	0.05
OCI Core	0.13	0.88	0.90	0.51
ECI Core	0.16	0.62	0.79	-0.02
Comms. (All Ties)	0.19	0.25	0.50	-0.15
Comms. (Business Ties)	0.15	0.25	0.48	-0.20

Table 4.7 Weighted Global Metrics Across 1-Mode `Ndrangheta Networks

	Weighted ALCC	*Weighted Degree Assortativity*
OCI (Full, CO)	0.19	0.36
ECI (Full, CO)	0.06	0.01
OCI (Full, NW)	0.11	0.00
ECI (Full, NW)	0.05	-0.09
OCI (Core, CO)	0.24	0.30
ECI (Core, CO)	0.08	-0.04
OCI (Core, NW)	0.12	0.00
ECI (Core, NW)	0.06	-0.10

failed to adequately capture the global structural characteristics of the ECI and communications networks.

4.4.5 Local Network Structure Comparisons

The observed dissimilarity in global metrics, however, obscures the similarity of the OCI network and ECI network at the local neighborhood-level. Tables 4.8 and 4.9 present the unweighted and weighted NetSimile results respectively. The tables present the overall NetSimile score for each network pair as well as the similarity for each of the seven features considered: (1) number of nodes in the egonet, (2) local clustering coefficient, (3) average neighbor degree, (4) average neighbor local clustering coefficient, (5) number of edges in the egonet, (6) the number of outgoing edges from the egonet, and (7) the number of neighbors of the egonet.

The unweighted Full OCI network and the unweighted Full ECI network had similar local neighborhood-level structural characteristics. The NetSimile scores for the unweighted Core and weighted Full OCI and ECI networks were marginally lower, but still suggestive of strong structural similarity. The OCI network did a worse job of capturing the neighborhood structures of the communication networks. The features with the weakest similarity tended to be (6) outgoing edges from egonet and (7) number of neighbors of the egonet for the unweighted NetSimile algorithm and (2) weighted local clustering coefficient and (4) average neighbor weighted local clustering coefficient for the weighted NetSimile algorithm.

To contextualize the magnitude of the NetSimile scores, Figure 4.1 presents the results of a heuristic rewiring test in-line with the approach used in Berlingerio et al. (2013). Synthetic networks were generated by taking one of the ECI or communications networks and rewiring a number of its edges; rewiring randomly took two edges in the observed network *i-j* and *k-l*,

Table 4.8 Unweighted NetSimile Results for Pairs of 1-Mode `Ndrangheta Networks

	Overall	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
OCI Full, ECI Full	0.76	0.78	0.73	0.87	0.68	0.76	0.78	0.76
OCI Core, ECI Core	0.68	0.76	0.74	0.86	0.65	0.71	0.54	0.53
OCI Full, Comms. (All Ties)	0.58	0.74	0.59	0.70	0.57	0.68	0.49	0.31
OCI Full, Comms. (Business Ties)	0.60	0.76	0.59	0.76	0.49	0.63	0.54	0.44
ECI Full, Comms. (All Ties)	0.68	0.85	0.68	0.80	0.51	0.90	0.51	0.49
ECI Full, Comms. (Business Ties)	0.72	0.92	0.70	0.83	0.61	0.85	0.58	0.58
Comms. (All Ties), Comms (Business Ties)	0.87	0.92	0.97	0.92	0.81	0.86	0.85	0.78

Table 4.9 Weighted NetSimile Results for Pairs of 1-Mode `Ndrangheta Networks

	Overall	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
OCI (Full, CO), ECI (Full, CO)	0.67	0.78	0.32	0.82	0.44	0.84	0.73	0.76
OCI (Full, NW), ECI (Full, NW)	0.69	0.78	0.53	0.68	0.53	0.86	0.72	0.76
OCI (Core, CO), ECI (Core, CO)	0.61	0.76	0.51	0.72	0.43	0.82	0.50	0.53
OCI (Core, NW), ECI (Core, NW)	0.64	0.76	0.70	0.59	0.63	0.81	0.46	0.53

removed these edges from the network, and added edges *i*-*k* and *j*-*l*. This rewiring procedure maintained the degree of each node in the network. The NetSimile score was then calculated to determine the similarity between the rewired network and the original observed network. For each share of edges to be rewired, 100 simulations were run and the average NetSimile score is reported in Figure 4.1. Horizontal dotted lines report the observed NetSimile score between the ECI and communications networks and the OCI network of the same characteristics – unweighted or weighted, count overlap or Newman projection, and full or core.

Figure 4.1a shows that the NetSimile score between the full unweighted OCI and ECI networks was similar to the score between the ECI network and simulated networks in which 15% of edges had been rewired; this increased to approximately 35% for the core ECI network. Using the degree-maintaining rewiring procedure did not yield NetSimile scores as low as the observed values for the two communications networks, suggesting there were material differences between these networks and the unweighted full OCI network. Figure 4.1b suggests that the NetSimile scores from comparing the weighted OCI and ECI networks were similar to those obtained by rewiring approximately 30% of edges for the full networks and approximately 55% of edges for the core networks. These simulations along with the observed NetSimile scores, while only illustrative, suggest that the OCI networks were best suited for representing the underlying local neighborhood structures of the ECI networks, especially the unweighted networks and the weighted full networks. On the other hand, the OCI network was a poor proxy for the local neighborhood structure of the communications networks.

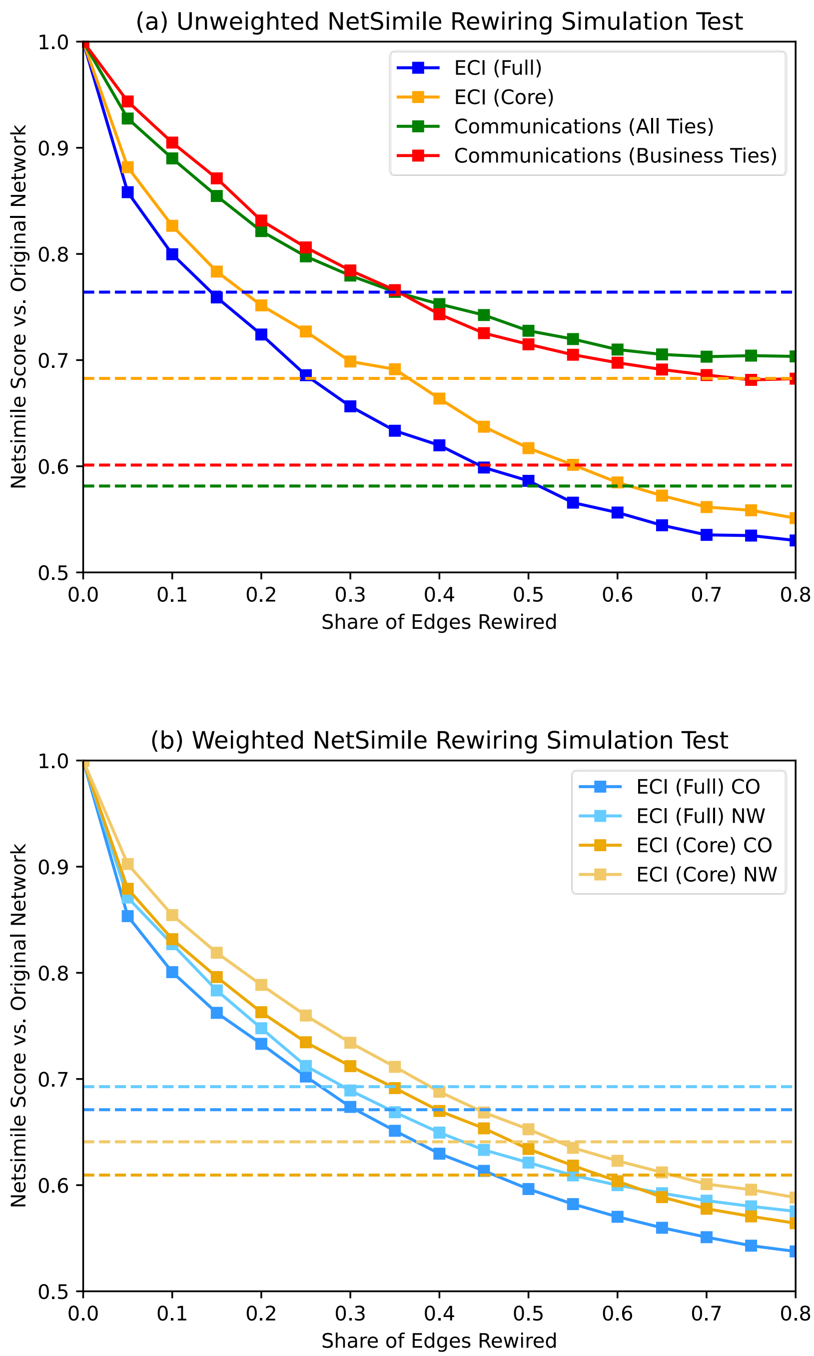


Figure 4.1. NetSimile scores between closed-source networks and rewired simulations for (a) the unweighted networks and (b) the weighted networks. Observed NetSimile scores between each network and its OCI equivalent in the dotted horizontal line.

4.5 Discussion

The findings reveal that network data from publicly-available business registers can be beneficial for investigating the structure of organized crime groups as they infiltrate legal businesses. These data, however, should be used in specific ways to ensure their validity. Based on a qualitative assessment of the strengths and weaknesses of both open-source and closed-source data, closed sources like judicial files and police data have the advantage of containing rich and hard-to-uncover information on criminal activities. On the other hand, business registers enable researchers to investigate a wider set of questions with more comprehensive and complete data. They contain complete information on the universe of relevant companies and actors that allows researchers to make boundary specification decisions that best fit their research questions rather than forcing these decisions to be made by law enforcement agencies and prosecutors who have different priorities. These registers also provide rich information on companies and corporate structures that may be incomplete in closed sources, as well as granular information on the temporal dimension of when actors form, maintain, and dissolve their affiliations to companies and data on how network structures evolve following law enforcement disruption. Commonly-used closed sources lack a comprehensive view of the granular data needed for more in-depth network analyses of organized crime group structure. Business registers also enable research in contexts where official data may be inaccessible or non-existent.

The overlap between the data sources was also compared. The OCI network contained over 70% of the actors in the ECI network and over 70% of the edges between these actors. However, the OCI network contained a smaller share of the nodes and edges in the two communication networks analyzed. While imperfect, the OCI network captured the majority of the network data from the ECI network when studying corporate interlock of organized crime group members. Further, 13% of the companies in the ECI network had erroneous or incomplete data based on the information contained in pre-trial detention order. This suggests that triangulation with business register data is necessary when using corporate interlock networks from closed sources to ensure the accuracy of the information.

While useful as a standalone data source for investigating questions relating to the overall structure of attempts to infiltrate legal markets and as a supplement to other sources, researchers must be wary when using network data from publicly-available business registers as a proxy for closed-source data. Despite the overlap in nodes and edges, the fact that covert ties were missing from the OCI network resulted in poor similarity on node- and network-level structural metrics – the piece of the picture missed by OSINT led to materially different results. The network data from the business register did not reliably produce similar results both for which actors were most central and for whole-network metrics, both for the communications networks and for the ECI network that included covert ties.

On the other hand, the OCI network was able to capture local neighborhood-level structures of the ECI network that included covert ties, even when node- and network-level structural metrics meaningfully diverged. The similarity between the networks was strongest when looking at the full and unweighted versions of the networks, with marginally weaker similarity when comparing the core or weighted versions of the networks. While capturing some structural features of the communications networks, the OCI network was less reliable overall for representing these networks.

The comparisons in the prior analysis represent an approximate lower bound on the performance of relational data from business registers as a proxy for closed-source data on corporate interlock and communications. The OCI network differed meaningfully from the ECI and communications networks on a range of metrics. In other cases with fewer or less impactful covert ties, business register data may work better as a proxy for node-level and global results; however, this cannot be known at the outset without the collection of closed-source data. As such, researchers should assume that network data from business registers will fail to act as a reliable proxy for closed-source data when looking at node- and network-level structural results unless they have specific evidence to the contrary. Despite these differences, however, the OCI and ECI networks exhibited similar local network structures.

These results have implications for academic research, as well as policy and policing. While business registers have been used to construct corporate interlock networks for non-infiltrated businesses, caution must be taken when using these same sources to investigate corporate interlock within organized crime groups. Whole-network metrics like degree centralization and clustering are likely to materially differ between corporate interlock networks based on business register data and police or judicial files even when representing the same underlying social phenomenon. Results on local structural configurations, on the other hand, are more likely to be reliable. Exponential random graph models (ERGMs) and stochastic actor-oriented models (SAOMs) model how actors form ties based on local network configurations (Bright et al., 2019; Lusher et al., 2012). These models may provide similar results when applied to corporate interlock networks based on business register data and closed-source data that include covert ties based on the observed similarities in local neighborhood structures. While further research is needed to understand and quantify the differences in the two data sources, the results in this paper suggest the use of ERGMs and SAOMs with business register data may be suitable for investigating tie formation processes in corporate interlock networks for organized crime groups, especially in contexts where closed-source data are not available.

Beyond academic relevance, law enforcement and policymakers can use business register data to better understand how organized crime groups infiltrate legal markets and how this can be combatted. The poor similarity in central nodes between the OCI network and the ECI and communications networks, as well as the fact that the business register data do not capture covert ties, means that this OSINT network data cannot and should not be used as evidence. The results in Section 4.4, however, suggest that business register data should be used to confirm the veracity of official company ownership information. Further, business register network data may be useful as a low-cost means of guiding early-stage law enforcement investigations. Connections between companies and the positioning of certain actors could provide a means of prioritizing surveillance and investigative operations, although further research is required to determine the efficacy of this approach.

4.6 Conclusion

Infiltration of legal businesses allows criminals to expand their influence and launder the proceeds of their illegal activities. While academic research has investigated aspects of this infiltration, the structure of organized crime groups as they found, buy, and partner with legitimate companies has not been studied. While publicly-available business registers present a potential data source for studying criminal networks in legal markets, their usability must be confirmed given concerns that business registers may not reliably represent underlying criminal structures since they are based on self-reported publicly-available information. This paper validated the reliability of network data from publicly-available business registers for investigating criminal network structures in legal businesses by comparing these data to more commonly-used closed-source data from a pre-trial detention order on the `Ndrangheta. A heuristic assessment suggested that business registers are suitable as a standalone data source for research focusing on the overall structure of infiltration rather than on just the organized crime group members, especially if the research design involves granular temporal network data or company characteristics. Results indicate that when closed sources like judicial files and police data on infiltration are available to researchers, network data should still be collected from publicly-available business registers to supplement these sources, which may be incomplete or contain erroneous information on company affiliations.

Researchers must be cautious, however, when using relational data from business registers as a proxy for closed sources. The OCI network did not reliably capture either node-level or global structural characteristics of networks derived from a pre-trial detention order. However, the data were able to capture local neighborhood-level structures of the ECI network, suggesting that business register network data is best suited for analyses that focus on local network structures. While business registers can provide a rich source of network data, researchers must be careful to avoid using the data to answer questions to which they are not suited.

This research contributes to the literature by testing and validating the usability of a new data source for investigating organized crime infiltration of legal markets. This is particularly important for contexts where infiltration has not been studied that may lack high-quality judicial or police data. Using the types of analyses for which the business register data appear to be well-suited, researchers can investigate topics like industry specialization, clustering and community structure, the role of brokers, and the differences in organized crime group structures in legal and illegal markets. More generally, these findings highlight the importance of checking the reliability of a new data source. OSINT and publicly-available data present an opportunity for academic research to investigate new groups, countries, and criminal activities. The reliability of these sources, however, cannot be assumed, especially when studying phenomena like organized crime groups that intentionally hide their actions from public view. Instead, it is necessary for researchers to determine which questions, research designs, methodologies, and analyses can best be served by using OSINT and which require harder-to-access closed-source data instead. These findings have implications for both academic researchers and practitioners interested in organized crime activities in legal markets. While data on this topic has previously been difficult-to-collect or unavailable, the accessibility of business registers means that a wider range of groups, contexts, and topics can be studied.

Using network data from publicly-available business register presents an opportunity to investigate the structure of organized crime groups as they infiltrate legal businesses and test hypotheses about these structures, keeping in mind the analyses for which these data are well-suited. Three further avenues for future research involve further validation of OSINT data. First, similar analyses should be conducted on other organized crime groups, countries, and closed sources to determine whether business register data are more or less reliable in other contexts. Second, other OSINT sources should explicitly be tested to determine whether they can reliably be used for academic research. This includes combining other OSINT sources with business register data. The results in this paper suggest that business register data fail to capture the network data or structural characteristics of communications networks from closed sources. Utilizing other OSINT sources like publicly-visible connections on social media, marriage records, and media reporting could provide a richer view of the relationships among organized crime group members and associates, although these sources must also be evaluated. Third, this analysis was limited in that it was based upon a set of companies that had been identified from a pre-trial detention order. Future research should investigate how OSINT can be collected regarding organized crime groups in a fully open process without using supplementary closed-source data, such as the use of graph mining techniques to identify infiltrated companies in the overall business register.

5

Proxy Targeting and Multiplex Resilience for Organized Crime Network Disruption

Abstract[17]

Understanding the structure and resilience of organized crime groups like mafias is critical for designing strategies to disrupt their operations. This paper proposes a new protocol – proxy targeting – for network disruption analyses. By separating the targeting of nodes and their removal from the network of interest, this approach allows researchers to consider multiple dimensions in a multiplex network simultaneously, as well as the sort of information available to law enforcement in the early stages of an investigation rather than assuming complete knowledge of the network. Multiplex network data is collected from a pre-trial detention notice and the Italian business register on an `Ndrangheta group as its members and associates sought to infiltrate legitimate businesses. Using the newly-developed proxy targeting approach, results suggest that open-source business registers can be an effective, low-cost, and easy-to-access

[17] A version of this paper has appeared in: Breuer, N. (2025). Proxy targeting and multiplex resilience for organized crime network disruption in corporate interlock and communications networks. *Social Network Analysis and Mining*, *15*(78), 1-20. https://doi.org/10.1007/s13278-025-01506-y

network data source for targeting surveillance and disruption of corporate interlock dimensions that represent which actors are affiliated to the same company together; however, these data are less successful for targeting disruption of the communication dimension between actors. Further, findings suggest that multiplexity can reinforce a network's resilience against disruption by providing fallback connectivity in the event that one dimension is destabilized. The proxy targeting protocol proposed here creates opportunities to answer new questions and to better understand how criminal networks are structured and how they can be disrupted.

5.1 Introduction

A key topic for the study of mafias and other organized crime groups is resilience – how vulnerable are these organizations to destabilization from internal or external forces like power struggles and law enforcement intervention? How capable are they of withstanding or adapting to disruption? What strategies can law enforcement use to best target these criminal groups? Due to data limitations, most studies on the resilience of mafias and other organized crime groups are based on simulations on observed pre-disruption networks: nodes in the observed network are iteratively removed according to a given targeting criterion – often based on the node's position in the observed network – and global metrics summarizing the connectivity and efficiency of the network are calculated and compared to their initial values (Chen et al., 2007). This approach, however, fails to account for multiplexity – multiple different kinds of connections between actors like friendship, kinship, and co-offending – and is of limited operational value to law enforcement because it utilizes information that is typically not available to police at the outset of an investigation prior to lengthy and costly physical and electronic surveillance (Musciotto & Miccichè, 2022).

This paper develops a new approach to network disruption and resilience analyses that addresses these limitations. Prior approaches – what I will refer to as 'native targeting' – use nodes' positions within an observed monoplex network to determine which should be removed from that same network. This paper proposes and applies a new approach in which targeting criteria are calculated on one observed dimension of a multiplex network and nodes are removed from a different dimension, for which connectivity metrics are then calculated. This approach is referred to as 'proxy targeting' because rather than being chosen for removal based on their position within the dimension being disrupted, nodes are instead targeted based on their position within a different dimension. Separating the targeting and disruption calculation functions in the analysis allows researchers to better consider and understand the different levels of resilience for different dimensions within a multiplex network simultaneously; similarly, this approach can allow law enforcement to determine whether easily-accessible information can be used to target surveillance on key actors whose arrest would most weaken the connectivity of the hidden underlying criminal network structure.

This paper tests this newly-proposed proxy targeting approach by studying early-stage investigative targeting strategies and multiplex resilience in a mafia network. Using network data from the Italian business register – Il Registro Imprese – and a pre-trial detention order detailing an `Ndrangheta group's attempts to infiltrate legal businesses in Italy and Germany, the resulting multiplex network contains official corporate ties, covert company affiliations, and criminal communications. Two main analyses evaluate the proxy targeting approach. First, business register data prove useful as a low-cost means of targeting surveillance of key actors in the early stages of law enforcement investigations for the corporate interlock dimension but are less effective for communication ties between mafiosi. Second, I investigate resilience in multiplex networks by simultaneously considering the indictment-based corporate interlock dimension and the communications dimensions using proxy targeting. Results suggest that the resilience of a given mafia or organized crime group can differ by the sort of network ties considered – a native targeting strategy that disrupts one dimension does not necessarily function well as a proxy targeting strategy to disrupt another dimension. Thus, future studies of resilience should move beyond only considering the disruption of one aspect of a mafia or organized crime group and instead look at resilience from a multiplex perspective. However, while multiplexity strengthens resilience, this effect weakens as more nodes are removed.

This research contributes to the existing literature in three principal ways. First, I extend disruption analysis from considering purely-illegal activities like drug trafficking to mafias and organized crime groups in legal markets, an area that leads to particular network structures that may be more or less resilient to disruption than structures focused on conducting illegal activities. Second, the newly-proposed approach broadens the focus of network resilience to include the effects of multiplexity by enabling researchers to consider multiple dimensions simultaneously rather than aggregating different layers into a single monoplex network or considering them separately. Third, the proxy targeting approach allows for the development and testing of new law enforcement targeting strategies that are based on the type of information that is available prior to the time-consuming process of collecting surveillance-based information.

The article is structured as follows. Section 5.2 reviews the relevant literature on network disruption and multiplexity in organized crime group networks. Section 5.3 lays out the data, methods, and limitations. Section 5.4 presents the results of the two main analyses. Section 5.5 discusses the results and their implications for academic research and operational practice, while Section 5.6 concludes.

5.2 Literature Review

The literature on the disruption of organized crime group networks has utilized two primary approaches. On the one hand, some studies analyze information on the structure of groups both before and after law enforcement interdiction or another disruption. For example, Berlusconi (2022) and Diviák et al. (2022) examine how a drug trafficking network and two terrorist networks adapt after key actors are arrested, both finding that the networks respond to external disruption by relying on triadic closure and pre-existing social ties; further, the drug trafficking network exhibited a decrease in centralization and longer path lengths (Berlusconi, 2022). Other studies have examined networks facing other forms of destabilization, like the seizure of drug shipments (Morselli & Petit, 2007) and violent external shocks (Airola & Bouchard, 2020). While enlightening, these analyses are based on the availability of network data on groups both before and after they are disrupted. This is rare, especially because most studies of organized crime group structure make use of police or judicial data which typically do not contain information about a group after law enforcement interdiction has occurred. This limits the cases and contexts in which organized crime group disruption can be studied and constrains the analysis to opportunistic research based on the availability of data.

The other strand of research has instead focused on artificial analyses. Rather than using observations both before and after a group has been destabilized, these studies instead use network data on undisrupted groups and simulate how the network would change as nodes or edges are removed as a proxy for law enforcement interdiction or external shocks. Most of these analyses are static – global connectivity metrics are calculated on the resulting network once the node is removed and potential post-disruption adaptations are not considered. Many static disruption analyses target nodes for removal based on structural targeting criteria that prioritize nodes based solely on their position in the observed network; targeting criteria include both well-known centrality measures like degree, betweenness, and closeness (Agreste et al., 2016; Cavallaro et al., 2020; Jahanpour & Chen, 2013; Morselli & Roy, 2008; Wood, 2017; Yeung et al., 2025) and newly-proposed measures of node prominence which are then validated using the disruption analysis (De Andrade et al., 2021; Jia et al., 2024; Musciotto & Miccichè, 2022; P. B. et al., 2023). More recently, researchers have sought to integrate a human capital targeting approach that prioritizes nodes for removal based on their role, knowledge, skills, and resources, either as a standalone approach or in conjunction with structural targeting approaches (Ficara, Curreri, Fiumara, De Meo, et al., 2022). While evidence of the efficacy of human capital targeting is mixed (De La Mora Tostado et al., 2024; Villani et al., 2019), two studies that directly compared human capital targeting approaches to structural targeting criteria found that structural criteria tended to outperform human capital approaches and that considering node attributes like role did not meaningfully increase the efficacy of centrality-based targeting (Bright et al., 2014; Ficara, Curreri, Fiumara, & De Meo, 2022).

These static analyses, however, fail to account for the adaptations organized crime groups undergo following a disruption observed in empirical post-disruption analyses. To account for this, a number of recent studies – based on disruption analyses using both structural and human capital targeting approaches – extend the analysis beyond the moment of disruption by allowing for network adaptation and recovery before calculating the global connectivity metrics used to evaluate the effectiveness of the disruption. The adaptation processes used include protocols that select another node in the network as a substitute in some or all of the removed node's edges (Bright et al., 2017; Duijn et al., 2014; Toledo et al., 2023), theoretically-driven network evolution modelling (Diviák, 2023), and agent-based modelling approaches based on observed case studies of criminal network recovery after law enforcement interdiction (Manzi & Calderoni, 2024b, 2024a). Based on a common goal of understanding criminal network vulnerability and designing more effective disruption strategies, progress in the field has been driven by new innovations in structural metrics and human capital approaches to more effectively identify key actors and developments in modelling the dynamics of post-disruption recovery to bring simulation-based results more in-line with empirical expectations and results.

This commonly-used approach, however, faces two primary limitations. First, it fails to account for multiplexity. Monoplex networks consist of only one layer of nodes and edges, with edges representing a single type of relationship or aggregating multiple connection types into a single combined tie. Multilayer networks, on the other hand, are collections of graphs – also called layers – and interconnections between those graphs (Boccaletti et al., 2014). Multiplex networks are a special case of multilayer networks in which each layer contains at least one shared node with another layer and the only inter-layer connections are between a node in one layer and its counterparts in other layers (Kivelä et al., 2014).[18] These layers allow for multiple dimensions of relationships between nodes. In the case of social networks, nodes may be connected by ties representing communication, collaboration, and kinship. These relationships can also exist in different contexts – in the case of mafias and other organized crime groups, for example, two members could be friends socially, work together on a criminal task, and hold shares in legitimate economic ventures together (Malm et al., 2010; C. M. Smith & Papachristos, 2016). Prior research suggests that organized crime groups are inherently multiplex organizations: their members exist in a web of interlocking relationships across different dimensions and contexts both with each other and with other external actors. For example, C. M. Smith & Papachristos (2016) look at the criminal, business, and personal networks connecting criminals in Prohibition-era Chicago associated with Al Capone; they find that while multiplex edges were rare, these connections were instrumental for both connecting criminals together and "[integrating] the underworld with legitimate society" (C. M. Smith & Papachristos, 2016, p. 663). Similarly, Diviák et al. (2019) find that pre-existing

[18] Some definitions of multiplex networks require that all layers contain the same set of nodes (F. Battiston et al., 2017; Boccaletti et al., 2014). The definition put forward by Kivelä et al. (2014) used in this paper relaxes this constraint.

social ties facilitated the development of communication and collaboration ties in a Czech corruption network, while Malm et al. (2010) compare the different types of relationships among Canadian criminals, finding that kinship and co-involvement in legitimate businesses act as the "'glue' that holds the larger criminal enterprise network together" (Malm et al., 2010, p. 68).

The findings of Malm et al. (2010) and C. M. Smith & Papachristos (2016) suggest that a small number of actors were key for connecting the different layers of the criminal systems to which they belonged. More generally, prior studies and reviews of the literature on multiplex networks suggest that multiplexity can make systems less resilient to disruption: "[T]he system responds globally to the damage inflicted in a single layer" (Boccaletti et al., 2014, p. 38; see also F. Battiston et al., 2017; Cellai & Bianconi, 2016). Other studies use the additional information provided by multiplexity to identify key actors; considering different types of ties has been used to identify actors who hold key roles in the Sicilian Cosa Nostra (Ficara et al., 2021) and to find actors who appear unimportant in the aggregated monoplex network but play an important role of enabling connectivity between networks of different tie types (Toledo et al., 2020). Despite the importance and utility of differentiating between distinct relationships and preliminary findings suggesting that disruption occurs differently in monoplex and multiplex networks, the standard approach to criminal network disruption has largely focused on monoplex networks of only one tie type. This approach fails to account for multiplexity and so cannot answer questions about the vulnerability and resilience of different aspects of a given network. For example, a monoplex approach cannot investigate whether the communication and collaboration dimensions of a multiplex network have different levels of resilience, whether they are vulnerable to different targeting approaches, and whether the disruption of the overall aggregated network may leave some key functions unaffected (Bright et al., 2015).

Second, the standard approach to network disruption simulations is of limited practical value for law enforcement because it presupposes that the researcher or analyst has complete data on the network structures of the organized crime group of interest. The standard monoplex approach typically focuses on networks comprised of communication or collaboration and co-offending ties, as these seem operationally relevant for understanding the criminal activities of an organized crime group. However, the datasets used to build these networks are typically based upon wiretapping, physical surveillance, and co-arrest data from police sources at the end of an investigation. These data, while helpful for understanding organized crime group structure from an academic perspective, are not available to law enforcement in the early stages of an investigation when targeting strategies for surveillance and disruption are most important (Manzi & Calderoni, 2024b; Musciotto & Miccichè, 2022). For example, a disruption strategy that identifies actors whose removal best destabilizes a communication network based on wiretap data cannot be used by law enforcement to identify actors it would be beneficial to wiretap: the data needed for the evaluation are only available after the point at which the evaluation would be useful. However, the standard monoplex approach must use post-investigation network data because these data offer the best representation of the group's operational structure and so are best suited for determining to what extent the network has been disrupted by a simulated attack. These analyses are constrained by the fact that the targeting strategies and simulated disruption are applied to the same base network.

Some studies have sought to address these issues. Regarding multiplexity, Bright et al. (2015) look at the degree and betweenness centrality of actors involved in a multiplex drug trafficking network to identify strategically-positioned nodes; while two nodes were highly central across the different dimensions, considering multiplexity uncovered other nodes who were important for connecting a specific aspect of the network like exchanging precursor chemicals or transferring money (Bright et al., 2015). Similarly, Malm et al. (2010) identify a set of actors in a Canadian criminal enterprise network that are high in betweenness centrality across a range of different tie types and suggest that these actors would be especially effective at disrupting the overall network if removed because they enable connectivity in multiple different dimensions. More recently, Toledo et al. (2023) incorporate information about nodes' connections in multiple dimensions of a multiplex Brazilian co-offending network into their targeting criteria for a network disruption analysis; they conduct the disruption analysis on the aggregated multiplex network, then separate the dimensions to understand how each is affected by the removal of nodes. Other approaches have sought to incorporate the type of information that law enforcement is likely to have available in the early stages of an investigation, like nodes' roles (Manzi & Calderoni, 2024b) and whether a node has connections to actors in other related organized crime groups (Musciotto & Miccichè, 2022).

Building on these prior approaches, the proxy targeting approach proposed in the following section aims to further develop our understanding of network disruption and resilience by decoupling the targeting function and the simulated disruption. Rather than using a node's position in an observed monoplex network to determine whether it should be removed from that same network, this new approach instead takes an observed multiplex network and uses its nodes' positioning in one dimension to target disruption of another dimension. This approach differs from traditional disruption analyses of organized crime groups by incorporating multiple dimensions of connections between nodes. Further, this protocol leverages the insights of Bright et al. (2015) and Malm et al. (2010) that certain actors can be key across multiple dimensions and extends their analyses by simulating what happens when these nodes are removed to test the prior expectations that removing these nodes would be especially damaging to a network's connectivity. The methodology of Toledo et al. (2023) is the most similar to the one proposed in this paper. Rather than using the aggregated monoplex network for both targeting and disruption before disaggregating into the individual dimensions to assess whether they were disrupted, the proxy targeting protocol instead allows for any dimension to be used as the targeting layer, extending the approach of Toledo et al. (2023). Finally, prior studies like Manzi & Calderoni (2024b) and Musciotto & Micciché (2022) have sought to understand whether more easily accessible information can be used to target the disruption of organized crime networks; however, information on actors' roles and their connections to other organized crime groups still require at least some difficult-to-collect police intelligence. The analyses described in the following section, on the other hand, test whether publicly-available network data from business registers can play a similar role.

This novel approach allows both academic researchers and law enforcement analysts to investigate a range of new theoretical and operational questions. For example, this protocol can be used to investigate convergence or divergence between the formal and informal structure of mafias by comparing the resiliency of the formally-defined role-based hierarchy of a mafia and the operational network of communication or cooperation between members: are these dimensions similarly vulnerable to disruption or do they diverge? Can the formal hierarchy provide a fallback for connectivity if the informal network is disrupted? What do these results reveal about the structure of mafias and how can the findings inform and test theoretical expectations? The proxy targeting approach can also be used to enrich role-based human capital targeting strategies by considering not just the level of a role like prior approaches, but also how that role connects to others within the mafia's hierarchy. More generally, sociologically and criminologically relevant questions about resilience for mafias, organized crime groups, and non-criminal organizations can be explored: does targeting nodes whose removal successfully disrupts one dimension of a multiplex network also disrupt its other dimensions? Does multiplexity increase resilience by reinforcing connectivity and creating redundancy or does the presence of key central actors in multiple dimensions create points of failure (Malm et al., 2010)? On the operational side, proxy targeting can allow for prioritization and surveillance targeting strategies to be developed and tested based on publicly-available network data that are accessible to law enforcement prior to expensive and time-consuming surveillance operations and that are richer than similar approaches in the existing literature based on role or connections to other groups. These multilevel and multiplex questions are gaps that prior approaches have struggled to investigate because they either focus on one network with one tie type or separately take each dimension of a multiplex network and treat them as monoplex. In the next sections, I lay out the proxy targeting approach and use it to answer questions about publicly-available network targeting and resilience in a multiplex mafia network.

5.3 Data and Methods

5.3.1 Data

The multiplex network used in the analyses in the next section is based on empirical data. The network represents a group of `Ndrangheta members and their associates active in Italy and Germany and involved in criminal activities ranging from racketeering and extortion to fraud, corruption, and attempted murder; beyond this, the group also sought to infiltrate legal markets by founding, investing in, and operating legitimate businesses in a range of industries including manufacturing, construction, hospitality, distribution, and agriculture. These companies were used for a range of illegal activities including laundering criminal proceeds, extorting other businesses, controlling legitimate markets, rigging public procurement contract auctions, and committing fraud. 169 of the group's members were arrested in January 2018.

Data were collected from two sources and were used to create three datasets. First, the pre-trial detention order relating to the case was manually coded for information on criminal communications. The document contains evidence of phone calls and in-person meetings between actors based on electronic and physical surveillance conducted throughout the investigation. In this communications dataset, actor-to-actor ties represent conversations and meetings between criminals where the topic of discussion was either directly about the infiltration and management of legitimate businesses or about other criminal activities. Actors and their associated communication ties were removed from the dataset if they were not charged with a crime in the pre-trial detention order, were not involved in any business-related conversations, and were not involved in company ownership based on evidence from the document; this filtering was done to remove irrelevant actors and conversations.

Second, information in the pre-trial detention order was used to construct a list of infiltrated companies. Companies were identified as infiltrated based on evidence in the indictment that they were owned or managed by mafiosi or their collaborators. This list was then used to create two corporate interlock datasets. For the first, historical ownership documents were obtained from the Italian business register, Il Registro Imprese, for each of the companies in the list. These documents contain the names of all actors who were officially involved in a company as a director, executive, or shareholder. Of the 165 companies in the list, ownership documents could not be obtained for 33 businesses for which the indictment did not include enough identifying information. Further, during the document collection process, three companies owned by actors charged in the pre-trial detention order were discovered and added to the list. Information on which actors were officially involved in which companies was extracted from these documents to create a business-register based corporate interlock dataset.

Third, the pre-trial detention order contained further information on company ownership and involvement. For each of the companies in the list, evidence from the indictment was used to identify connections between actors mentioned in the document and companies in the list. Unlike the business register, the pre-trial detention order did not mention all actors who were involved in each company; instead, a subset of actors deemed relevant to the case were included. This indictment-based corporate interlock dataset includes two types of ties. Similar to the business register-based dataset, there are official ties between actors and the companies in which they are an executive or shareholder. Comparison against the business register revealed a number of cases in which the official ownership information in the indictment was incorrect; in these cases, the indictment-based corporate interlock dataset was amended. Beyond official ownership, the pre-trial detention order also contained information on covert company involvement; in some cases, actors – especially mafiosi who feared that their actions were being monitored by the police – would secretly invest money into a company through a ‘clean’ accomplice or would maintain *de facto* control of a company without being officially affiliated to the business in a way that would show up in the business register. Information on which actors were covertly involved in which businesses was also coded. Based on the pre-trial detention order, actors who were officially or covertly affiliated to a company were the ones involved in its management and the execution of both legal and illegal activities through the business, with actors who were affiliated to the same company cooperating to run it. As such, these corporate interlock ties represent the main collaborative relationships as the group infiltrated legitimate markets.

These three datasets were used to create a multiplex network with four dimensions; Table 5.1 presents information and descriptive statistics for these dimensions and the aggregate monoplex network created by flattening the dimensions to a single layer. All four dimensions only include actors as nodes and are unweighted.

CI_O is based on the business register-based corporate interlock dataset. This dimension is comprised of official corporate interlock connections, where edges represent that two actors are involved in the same business in an official capacity based on the business register documents. All actors mentioned in the business register documents are included in this dimension. This dimension represents the sort of information that law enforcement could utilize prior to conducting surveillance or formal investigations.

CI_E is a dimension that represents the police intelligence-based view of corporate interlock. This dimension incorporates the view of covert affiliations provided by the pre-trial detention order. Edges occur between actors that are associated with the same business either officially or covertly based on the business register documents and the information in the pre-trial detention order. CI_E only contains nodes reported as relevant in the indictment to provide a more accurate view of the actors involved in legal-market infiltration by the `Ndrangheta group; nodes excluded from the order but present in the business register documents are not included in this dimension, while they are included in the CI_O dimension.

Table 5.1 `Ndrangheta Multiplex Network Information

		Dimensions			
	Aggregate Monoplex Network	CI_O	CI_E	$Comms_{AT}$	$Comms_{BT}$
Nodes	363	166	144	210	187
Edges	1,479	697	446	598	430
Descriptive Statistics					
Density	0.02	0.05	0.04	0.03	0.02
Average Degree	8.15	8.40	6.19	5.70	4.60
Number of Components	1	12	8	1	1
Degree Centralization	0.15	0.15	0.14	0.19	0.15
Degree Assortativity	-0.07	0.43	-0.07	-0.15	-0.21
Transitivity	0.50	0.88	0.64	0.25	0.25
Avg. Local Clustering Coefficient	0.70	0.92	0.76	0.51	0.48
Shared Nodes					
CI_O	-	166	73	36	34
CI_E	-	-	144	79	75
$Comms_{AT}$	-	-	-	210	187
$Comms_{BT}$	-	-	-	-	187

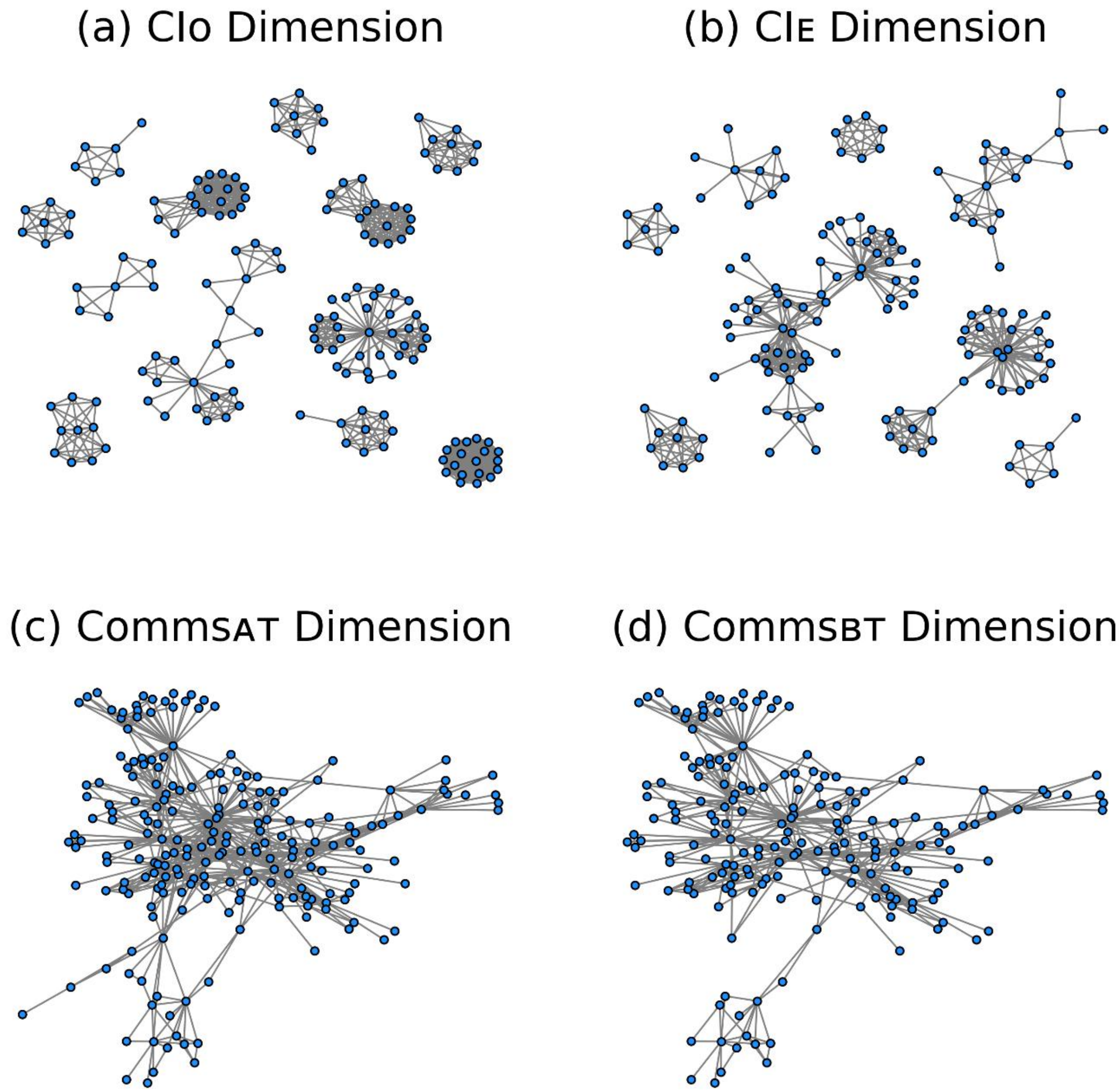


Figure 5.1. Network visualizations for the (a) CI_O dimension, (b), CI_E dimension, (c), $Comms_{AT}$ dimension, and (d) $Comms_{BT}$ dimension.

Finally, the network includes two communication dimensions: Comms_{AT} includes ties relating to any criminally-relevant communications, while Comms_{BT} only contains communication ties about the infiltration and operation of legitimate businesses. For the purpose of the following analyses, each dimension has been adjusted to remove components of five or fewer nodes; these small components are unlikely to contain relevant nodes based on the targeting metrics described below and could have unintended impacts on the performance metrics used. In the analyses described in Section 5.3.2, each dimension is treated as a separate network, although a multiplex graph object based on a supra-adjacency matrix could also be used (Kivelä et al., 2014). The four dimensions are visualized in Figure 5.1.

5.3.2 Methods

Figure 5.2 presents a schematic view of the native and proxy targeting approaches. The native targeting protocol that has commonly been used in prior studies typically starts by defining some targeting criterion or metric that is used to rank nodes in an observed monoplex network. As part of a disruption analysis, the top-ranking node or set of nodes is then removed from that observed network, at which point a set of performance metrics is calculated on the new subgraph from which the nodes have been removed; these values are then compared against the values of the performance metrics calculated on the original observed network. This protocol is applied iteratively for a range of node removal counts to observe how the network changes. Often, the removal happens sequentially – for a node removal count of five, the targeting metric is calculated on the observed network and the top-ranking node is removed to create a subgraph; the targeting metric is then calculated on the subgraph – accounting for changes induced by the removal of the first node – and the new top-ranking node is identified; and so on until five nodes have been removed. For the proxy targeting approach used here,

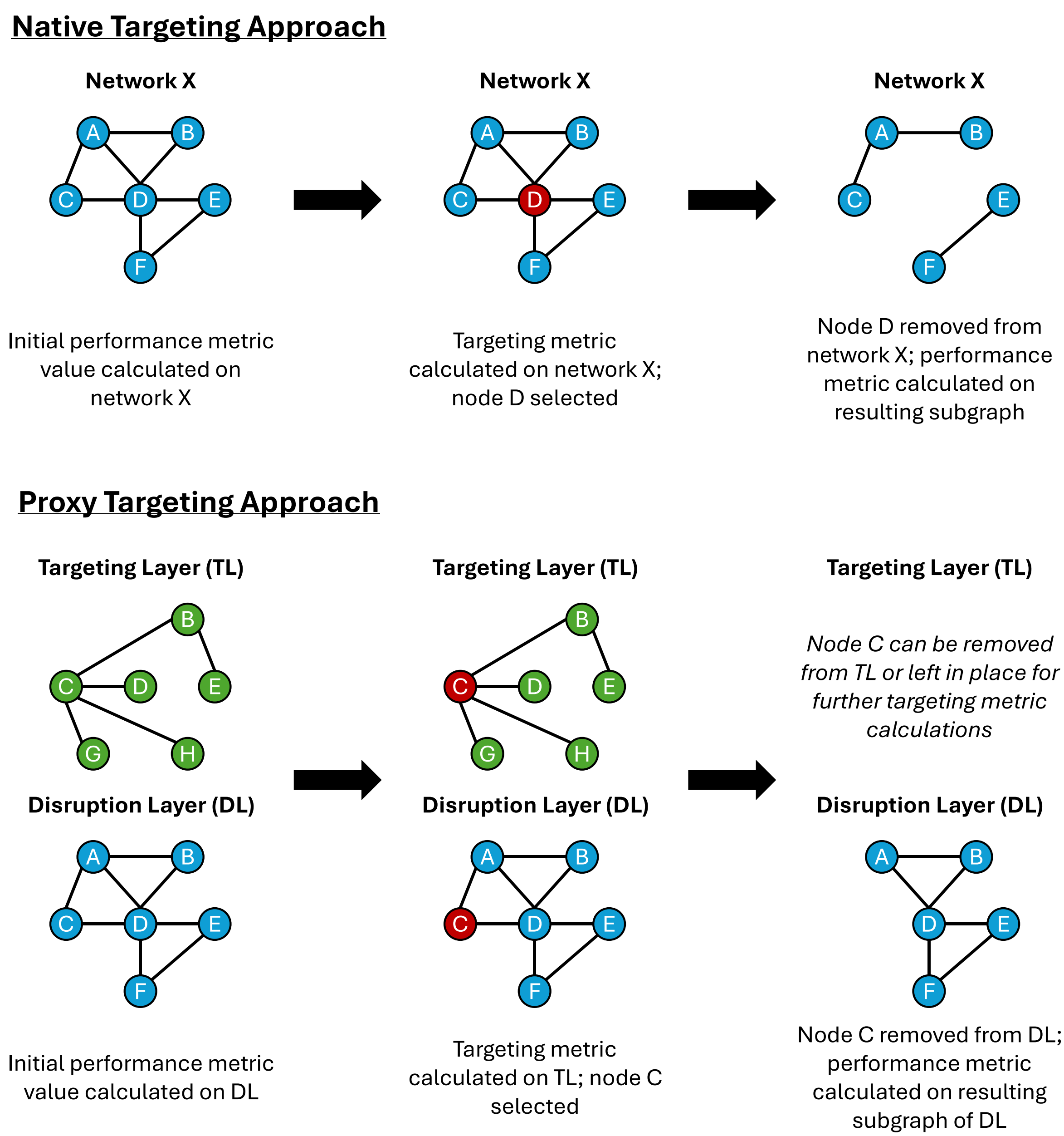


Figure 5.2. Schematic of the native targeting and proxy targeting approaches to network disruption analyses.

there are instead two different monoplex networks that contain overlapping nodes or two dimensions within a multiplex network.[19] In a disruption simulation, one dimension is labelled the targeting layer and the other is labelled the disruption layer. The targeting metric is calculated for nodes in the targeting layer based on that layer's edges; the top-ranking set of nodes is then removed from the disruption layer and the performance metrics are calculated based upon the new subgraph of the disruption layer. The targeting layer can either remain unchanged or the nodes can be removed sequentially to create a subgraph on which the targeting metric can be recalculated at each step; in the following analyses, the latter approach is used. In cases where the targeting and disruption layers are not comprised of the same set of nodes and only some nodes overlap between them, the researcher can choose whether to only consider nodes that are present in both layers or, when a node in the targeting layer that is not in the disruption layer is chosen for removal, that step can be counted without a node being removed from the disruption layer; the former choice is made in the following analyses as discussed in Section 5.3.3.

Four main analyses are conducted using the multiplex `Ndrangheta network. First, I tested the cohesion of the dimensions using both greedy and proxy targeting approaches. Similar to the Key Player Problem-Positive outlined in Borgatti (2006), this analysis aims to identify a set of nodes of a given size that are connected to the most other nodes in the network within some distance *m*. Here, this analysis is used to simulate how law enforcement might prioritize surveillance targets – what strategies can be used to determine which nodes should

[19] The proxy targeting approach can accommodate more than two networks or dimensions, with one network or dimension used for calculating the targeting metric and a number of other networks or dimensions used to calculate performance metrics. Based on the multiplex network used in the analyses within this paper and for the sake of clarity, the example only describes two networks or dimensions.

be wiretapped, physically surveilled, or investigated to provide coverage of the most other nodes in the network? For a given set size *n* and distance *m*, the greedy algorithm iterates over all permutations of *n* nodes and calculates the share of other nodes in the network that are at most *m* edges away from a node in the set and returns the set that reaches to most other nodes, thus representing the targeting set with the highest possible cohesion given the overall network structure; the proxy targeting approach, on the other hand, determines nodes for the set based on a targeting metric calculated on one layer and then calculates the share of other nodes in another layer that are within the distance *m* from those nodes in that layer. In Section 5.4, the analysis uses $n = 1$ and $n = 2$ for set sizes and a distance of $m = 2$, based on Friedkin (1983)'s horizon of observability, which suggests that in communication networks, actors are aware of the activities and performance of their alters – other actors with whom they are directly connected – and the neighbors of those alters.

I then carried out three disruption analyses. First, a native targeting disruption analysis was conducted using each of the three indictment-based dimensions as separate networks to investigate the resilience of each layer across a range of previously-used targeting criteria and performance metrics. Second, the CI_O dimension was used as a targeting layer while the other three dimensions were treated as disruption layers to determine whether the publicly-available business register network could be used by law enforcement to prioritize high-impact actors in the early stages of an investigation. Third, I investigated whether disrupting one dimension of the multiplex network also destabilizes other dimensions at the same time and whether multiplexity provides reinforcement and additional resilience to node removal. Proxy targeting disruption simulations used the police intelligence-based CI_E, $Comms_{AT}$, and $Comms_{BT}$ dimensions as both targeting and disruption layers in various combinations.

Across these analyses, degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality were used as targeting metrics in-line with prior studies (Agreste et al., 2016; Cavallaro et al., 2020; Yeung et al., 2025). Beyond the more commonly-used measures, the Alternative Path Centrality metric (APC) was also used; this centrality measure is similar to betweenness centrality, but accounts for both a node's centrality in the flow of information through a network and whether the network contains other equally-efficient paths for information flow if that node is removed (Shavitt & Singer, 2007). In cases where multiple nodes have the same targeting metric value, one of the nodes is chosen at random, determined by a random seed; to account for the variability introduced by these cases, 100 simulations were run and the results were averaged.[20] Four measures are used as performance metrics in the disruption analyses:

1. BF: Borgatti fragmentation based on the measure proposed in Borgatti (2006).
2. ALE: Average local efficiency, equal to the average of 1 minus the Borgatti fragmentation for the ego network of each node in the network.
3. LCC Size: The share of nodes in the largest connected component .
4. 3 LCC Size: The share of nodes in the three largest connected components to account for the greater number of components in the CI_O and CI_E dimensions.

An increase in Borgatti fragmentation suggests that the network has been disrupted and made less efficient, while disruption is signaled by decreases in the other three performance metrics. For the analyses using the proxy targeting approach, the effect size is calculated as the difference between the performance metric for the original layer and the average performance metric for the subgraph after node removal. Results are reported as the effect size achieved by proxy targeting divided by the effect size achieved by native targeting. A value of 1 suggests

[20] For the analyses presented in Sections 5.4.2-5.4.4, the integers between 1 and 100 were used as random seeds.

that native and proxy targeting on a given disruption layer for a given number of nodes removed were equally effective at changing the performance metric; a value below 1 suggests that the proxy targeting approach was less effective at changing the performance metric than the native targeting approach, while a value above 1 means the proxy targeting approach was more effective. This comparative approach was chosen to contextualize the proxy targeting results against what might be expected by natively removing the same number of nodes that are central in the disruption layer. For the native and proxy targeting approaches, one targeting metric was chosen for each approach for comparison; the targeting metric that had the largest average effect size for node removal amounts between 1 and 30 was chosen, with the choice occurring for each performance metric.

5.3.3 Limitations

While the analyses presented in Section 5.4 face some limitations based on the simulation choices made in the research design, the proxy targeting approach is flexible and can accommodate various modelling parameters and setups as dictated by different research questions and research designs. First, the proxy targeting protocol used in this paper only considers nodes that are present in both the targeting layer and the disruption layer – once the targeting metric is calculated, the top-ranking nodes of the set that belong to both layers are chosen and non-shared nodes are not considered. An alternative modelling choice would be to consider non-shared nodes – if a node in the targeting layer is selected by the targeting criterion but is not present in the disruption layer, it could take up a spot in the removal set but not impact the disruption layer. The shared node constraint was chosen to align with the sort of information that may be available to law enforcement at the beginning of an investigation – while the full network structure is not accessible, police will likely have knowledge of a subset of relevant actors of interest even prior to material electronic and physical surveillance; this knowledge is needed to make a determination that an investigation against a group or set of actors is warranted.

Second, the disruption analyses only consider structural targeting metrics and do not test the efficacy of human capital targeting or mixed approaches; centrality-based metrics are used rather than identifying nodes for removal based on their role, resources, or skills. These non-structural targeting criteria can be integrated with a proxy targeting approach, as can other structural metrics. Further, prior studies using a native targeting approach have suggested that structural targeting metrics generally perform as well or better than human capital measures (Bright et al., 2014, 2017; Ficara, Curreri, Fiumara, & De Meo, 2022).

Finally, the disruption analyses presented in Section 5.4 are static rather than dynamic: performance metrics are calculated on the disruption layer without allowing for post-disruption adaptation. Network adaptation can be used alongside the proxy targeting protocol to run dynamic proxy targeting disruption analyses. In this paper, the static approach was chosen for three reasons: (1) this allows the results to be compared to prior findings in the literature, which has predominantly used static analyses; (2) any adaptation protocol used would be assumption-based rather than rooted in empirical observation, limiting the realism and interpretability of the results (Diviák, 2023); and (3) pre-adaptation disruption results can still be useful for understanding vulnerability and resilience. A greater deviation from baseline performance metric values relating to connectivity and efficiency reduces a group's ability to respond and adapt to a destabilizing shock like arrests and likely requires a more drastic adaptation to regain pre-disruption structural characteristics, which can expose the network to further disruption (Duijn et al., 2014, p. 1).

Table 5.2 Greedy and Proxy Targeting Cohesion Results – n = 1, m = 2

		Proxy Targeting					
	Greedy Targeting	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*APC*	*Best Proxy Targeting as Share of Greedy Targeting*
CI_O	0.20	--	--	--	--	--	--
CI_E	0.36	0.22	0.22	0.22	0.06	0.22	61%
$Comms_{AT}$	0.78	0.40	0.40	0.40	0.06	0.40	51%
$Comms_{BT}$	0.61	0.33	0.33	0.33	0.05	0.33	54%

Table 5.3 Greedy and Proxy Targeting Cohesion Results – n = 2, m = 2

		Proxy Targeting					
	Greedy Targeting	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*APC*	*Best Proxy Targeting as Share of Greedy Targeting*
CI_O	0.31	--	--	--	--	--	--
CI_E	0.57	0.27	0.37	0.21	0.05	0.37	65%
$Comms_{AT}$	0.86	0.45	0.53	0.39	0.22	0.53	62%
$Comms_{BT}$	0.76	0.38	0.48	0.38	0.23	0.48	63%

5.4 Results

5.4.1 Cohesion Analysis Results

Tables 5.2 and 5.3 report results from the cohesion analyses with targeting set sizes of 1 and 2 nodes, respectively. In these analyses, the business register-based CI_O dimension was used for targeting on the other three dimensions; the values in the tables range from 0 to 1 and represent the share of non-targeted nodes in the network that are within a path distance $m = 2$ from the targeted nodes. For versions of the analysis where only one node was selected, all targeting metrics except for eigenvector centrality performed equally well; for a targeting set size of two, betweenness centrality and APC perform equally well and outperform the other structural targeting metrics. Focusing on the results in Table 5.3, the best proxy targeting results – based on betweenness centrality and APC – were 60-65% as effective as greedy targeting. For example, in the police intelligence-based corporate interlock dimension CI_E, greedy targeting yielded a set of two nodes from which 57% of the other nodes in the dimension could be reached within two steps, representing the best possible value for $n = 2$ and $m = 2$; the set of two nodes identified using the proxy targeting approach, on the other hand, could reach 37% of the other nodes in the CI_E dimension. A greater share of nodes in the police intelligence-based communication dimensions $Comms_{AT}$ and $Comms_{BT}$ were reachable based on greedy targeting – 86% and 76% respectively – and proxy targeting – 53% and 48% respectively.

5.4.2 Native Targeting Disruption Analysis Results

Figures 5.3-5.5 present the native disruption analysis results for the three police intelligence-based dimensions. In the CI_E dimension, the majority of the reduction in LCC size and size of the three largest components and of the increase in the Borgatti fragmentation occurred by the

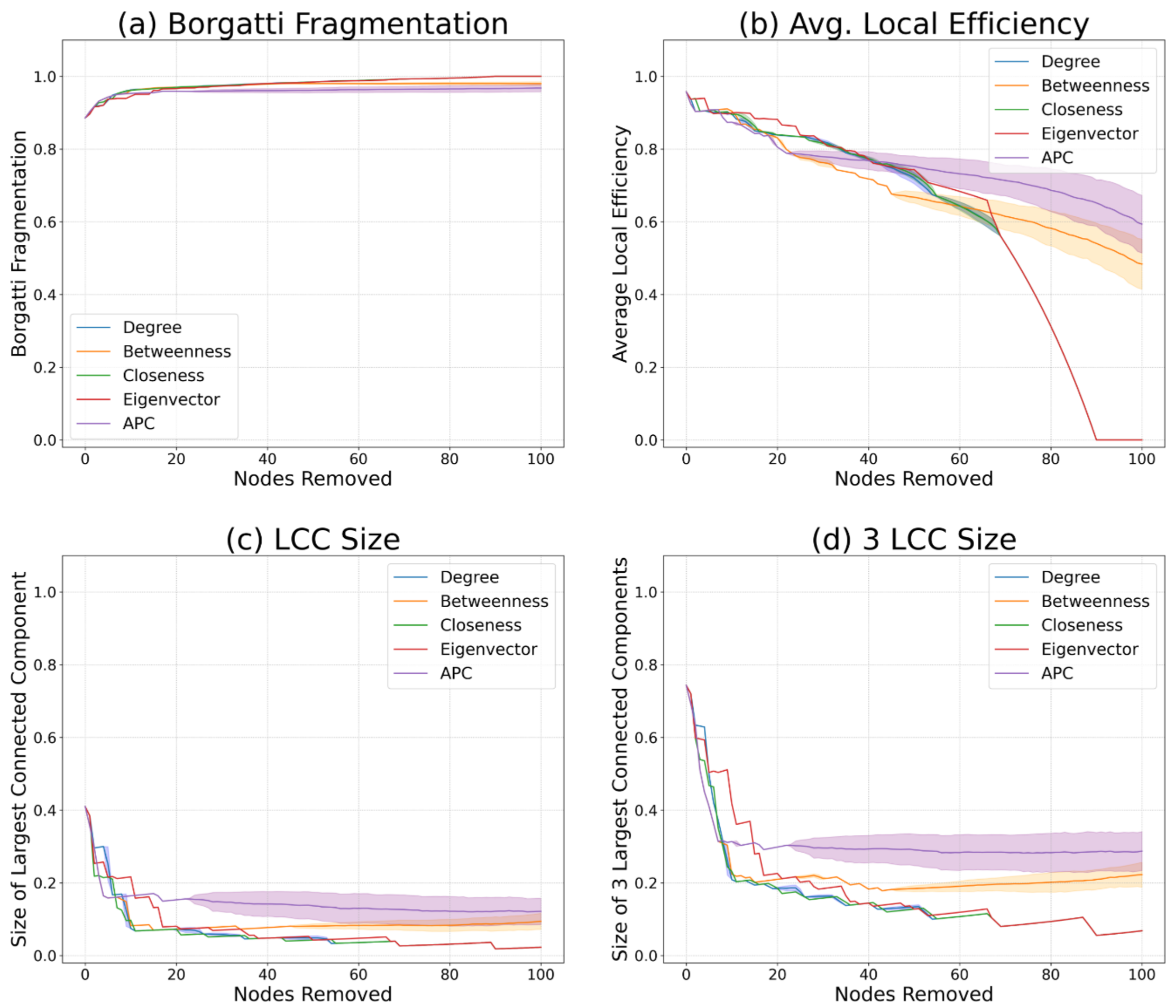


Figure 5.3. Native disruption plots for the CI_E dimension for (a) Borgatti fragmentation, (b) average local efficiency, (c) LCC size, and (d) 3 LCC size. The solid line represents the mean results across the 100 repeated simulations, while the shaded region indicates one standard deviation above and below the mean.

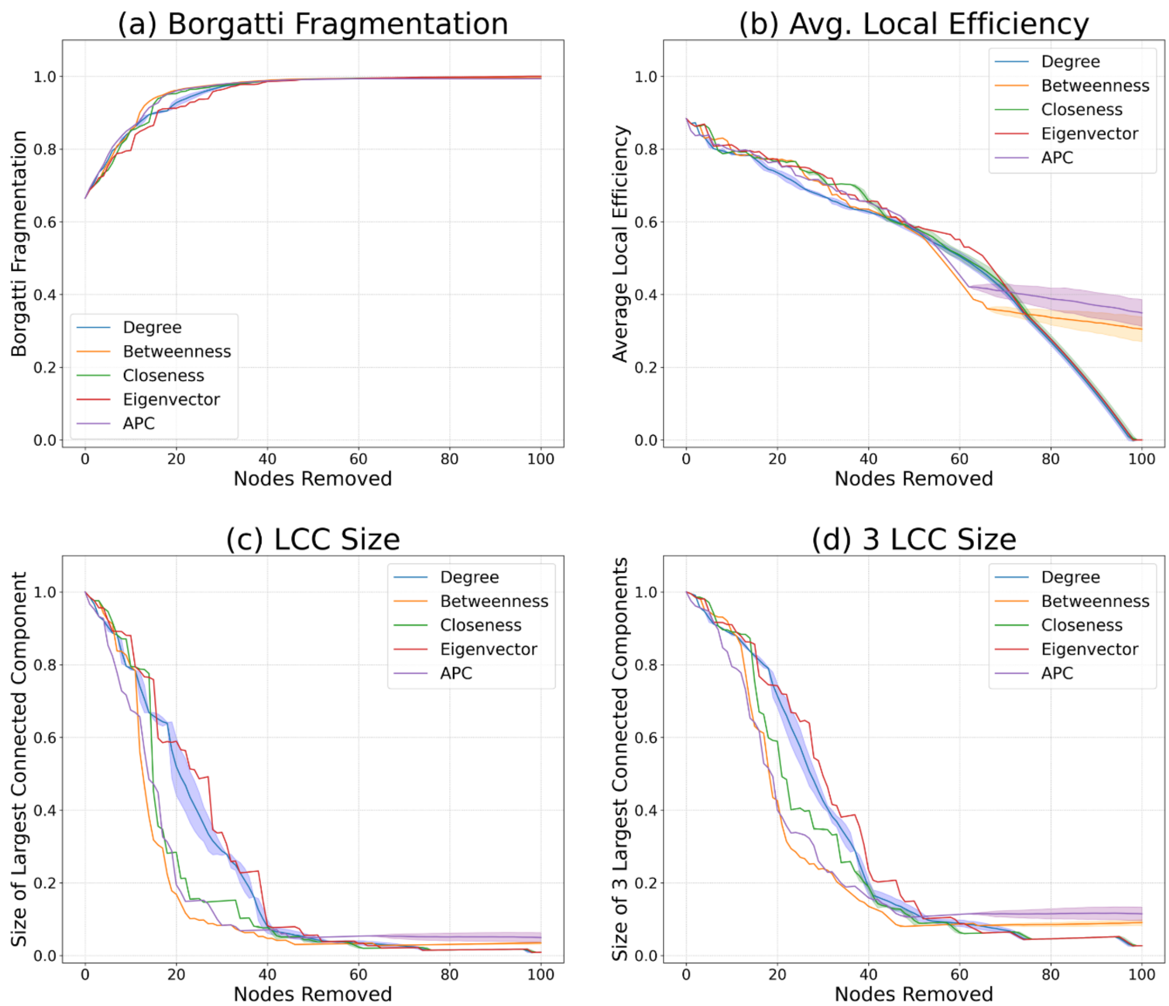


Figure 5.4. Native disruption plots for the Comms$_{AT}$ dimension for (a) Borgatti fragmentation, (b) average local efficiency, (c) LCC size, and (d) 3 LCC size. The solid line represents the mean results across the 100 repeated simulations, while the shaded region indicates one standard deviation above and below the mean.

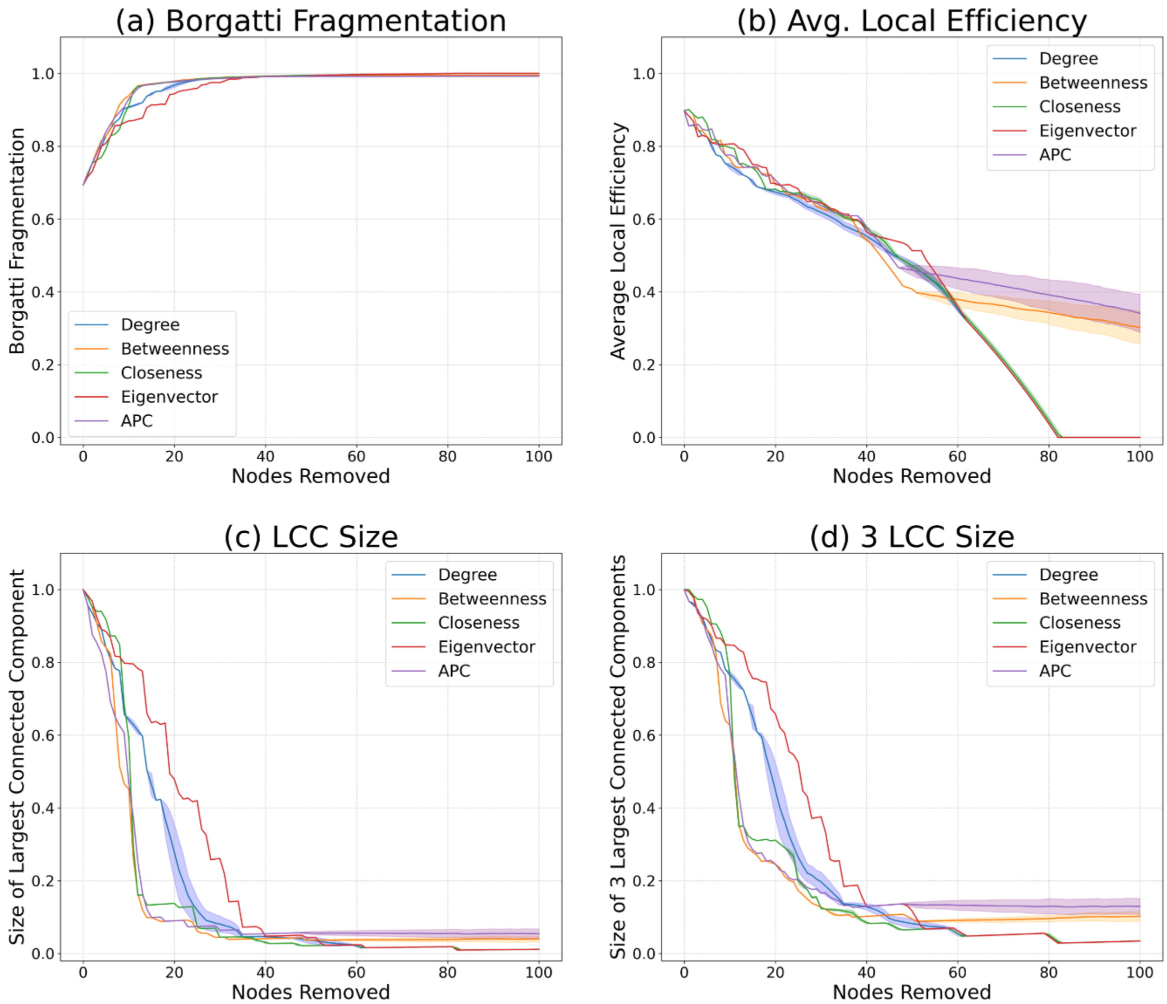


Figure 5.5. Native disruption plots for the Comms$_{BT}$ dimension for (a) Borgatti fragmentation, (b) average local efficiency, (c) LCC size, and (d) 3 LCC size. The solid line represents the mean results across the 100 repeated simulations, while the shaded region indicates one standard deviation above and below the mean.

time 10 nodes had been removed, visible as a flattening of the curves; the same can be seen for the Comms_{AT} and Comms_{BT} dimensions, typically after 20-25 nodes had been removed. The plots for average local efficiency exhibit more gradual and consistent declines as more nodes are removed; the rate of decline, however, is steeper for the communication dimensions rather than the corporate interlock dimension. These native targeting disruption results are used as the baseline against which the proxy targeting results in Sections 5.4.3 and 5.4.4 are compared.

5.4.3 Proxy Targeting Disruption Analysis Results – CI_O Targeting

To investigate whether publicly-available business register network data could be useful for targeting mafia networks in legal markets, a series of proxy targeting analyses were conducted in which the CI_O dimension was used as the targeting layer and the CI_E, Comms_{AT}, and Comms_{BT} dimensions were used as the disruption layers. Figures 5.6 and 5.7 present heatmaps in which the values represent the effect size of the proxy targeting divided by the effect size of the native targeting; a value of 1 means that both the native and proxy targeting had the same effect size, while a value less than 1 means that the native targeting approach outperformed the proxy targeting approach and a value greater than 1 means that proxy targeting outperformed the native targeting.

Figures 5.6 and 5.7 reveal that proxy targeting for disruption simulations using the CI_O dimension was materially more effective on the CI_E dimension than on either of the communication dimensions Comms_{AT} or Comms_{BT}. In Figure 5.6, removing one node identified through proxy targeting using the CI_O dimension is equally as effective as native targeting for the CI_E dimension at increasing Borgatti fragmentation and decreasing average local efficiency and more than twice as effective at reducing the share of nodes in the largest connected component and three largest connected components. For the first five nodes

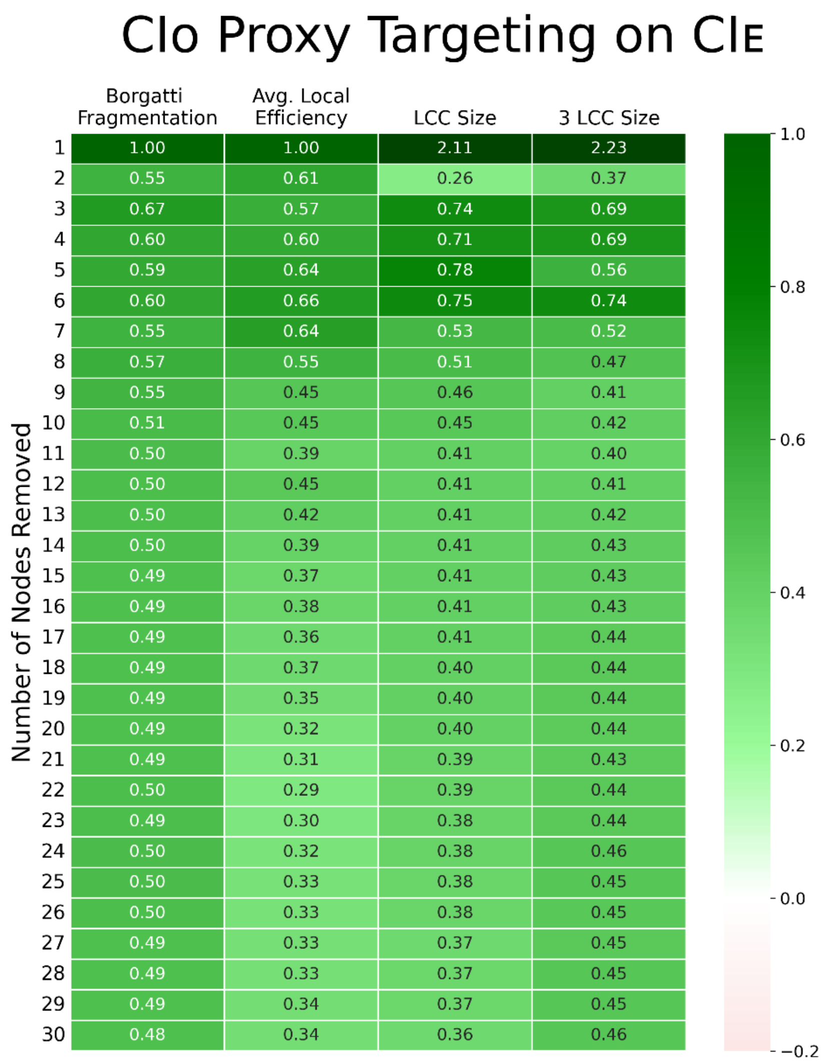


Figure 5.6. Heatmap for the results of CI_O proxy targeting on the CI_E dimension.

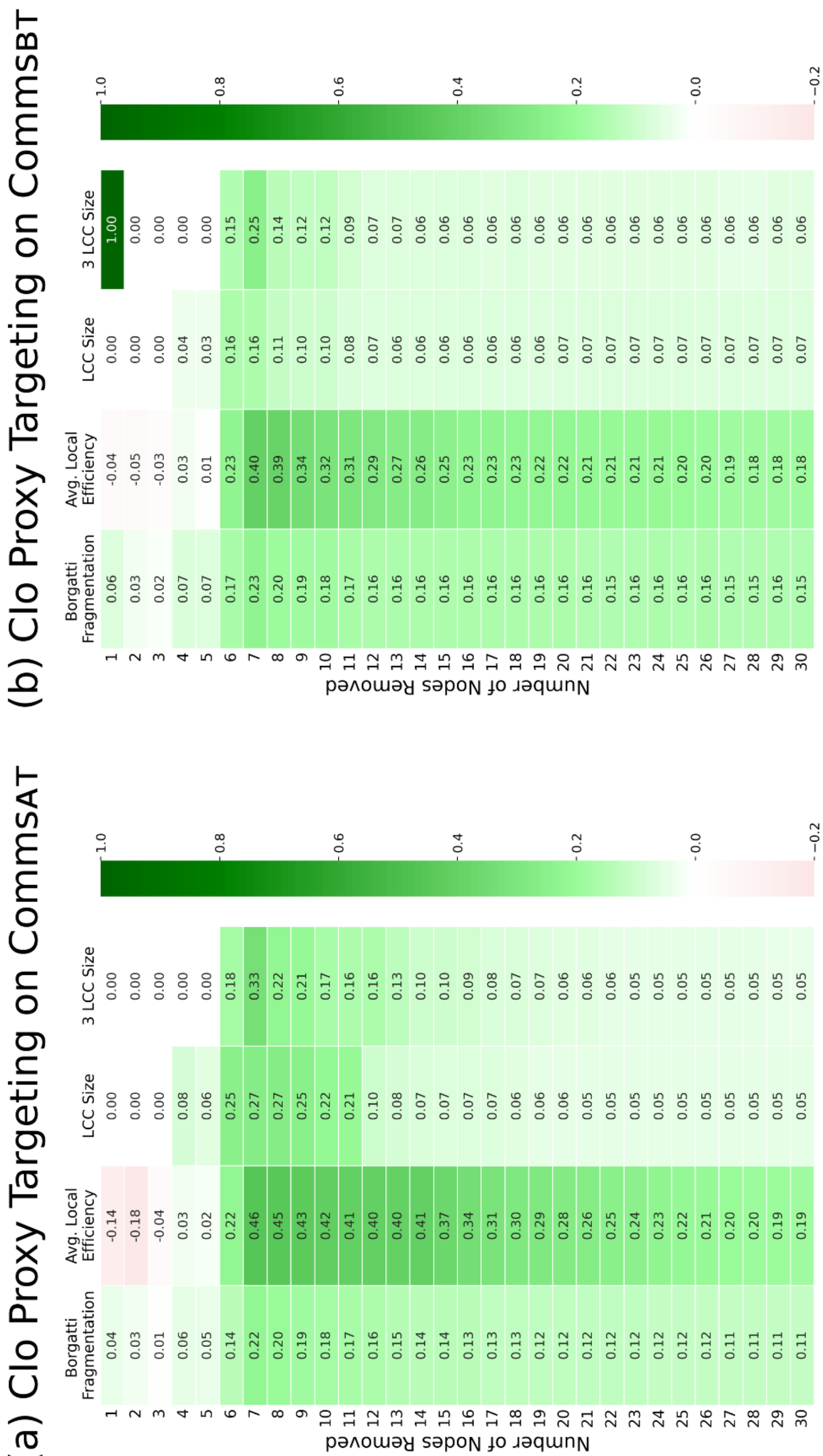


Figure 5.7. Heatmaps for the results of CIo proxy targeting on the Comms$_{AT}$ (left) and Comms$_{BT}$ (right) dimensions.

removed, proxy targeting is generally 55-80% as effective as native targeting; this drops to 30-50% effectiveness as the number of nodes removed increases.

On the other hand, proxy targeting using the CI_O dimension was not effective at disrupting the communication dimensions. Effectiveness for the first five nodes removed is below 10%, with one exception of 100% effectiveness at reducing the share of nodes in the three largest connected components when removing one node from the $Comms_{BT}$ dimension. In fact, proxy targeting using the CI_O dimension actually marginally increased the connectivity of the communication dimensions based on the average local efficiency for the first three nodes removed. For greater numbers of nodes removed, effectiveness increases but is still low; proxy targeting based on the CI_O dimension is best at reducing the average local efficiency of the communication dimensions, but only reaches a maximum of approximately 45% as much effect size compared to the disruption caused by native targeting for the same number of nodes removed.

5.4.4 Proxy Targeting Disruption Analysis Results – Multiplex Resilience

The final set of analyses – presented in Figures 5.8 and 5.9 – aim to investigate the multiplex resilience of the `Ndrangheta group and its associates: does a disruptive approach that destabilizes one dimension similarly impact other dimensions or does multiplexity reinforce the network against external shocks like node removal? The heatmaps presented here are similar to those in Section 5.4.3: Figure 5.8 shows the results from a disruption analysis where the $Comms_{AT}$ and $Comms_{BT}$ dimensions are the targeting layer and the CI_E dimension is the disruption layer, while Figure 5.9 treats the CI_E dimension as the targeting layer and the communication dimensions as disruption layers.

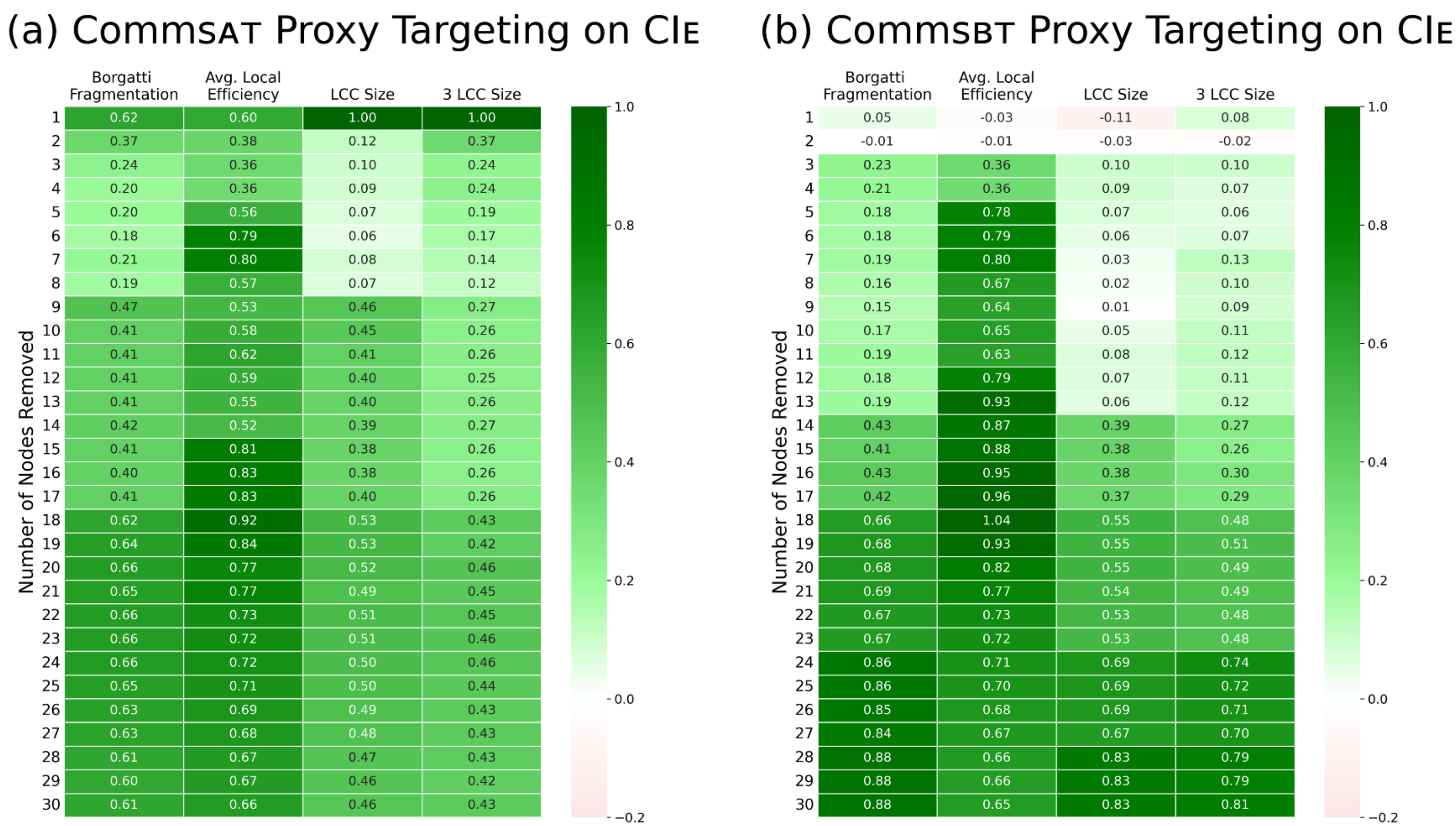


Figure 5.8. Heatmaps for the results of Comms_{AT} proxy targeting (left) and Comms_{BT} proxy targeting (right) on the CI_E dimension.

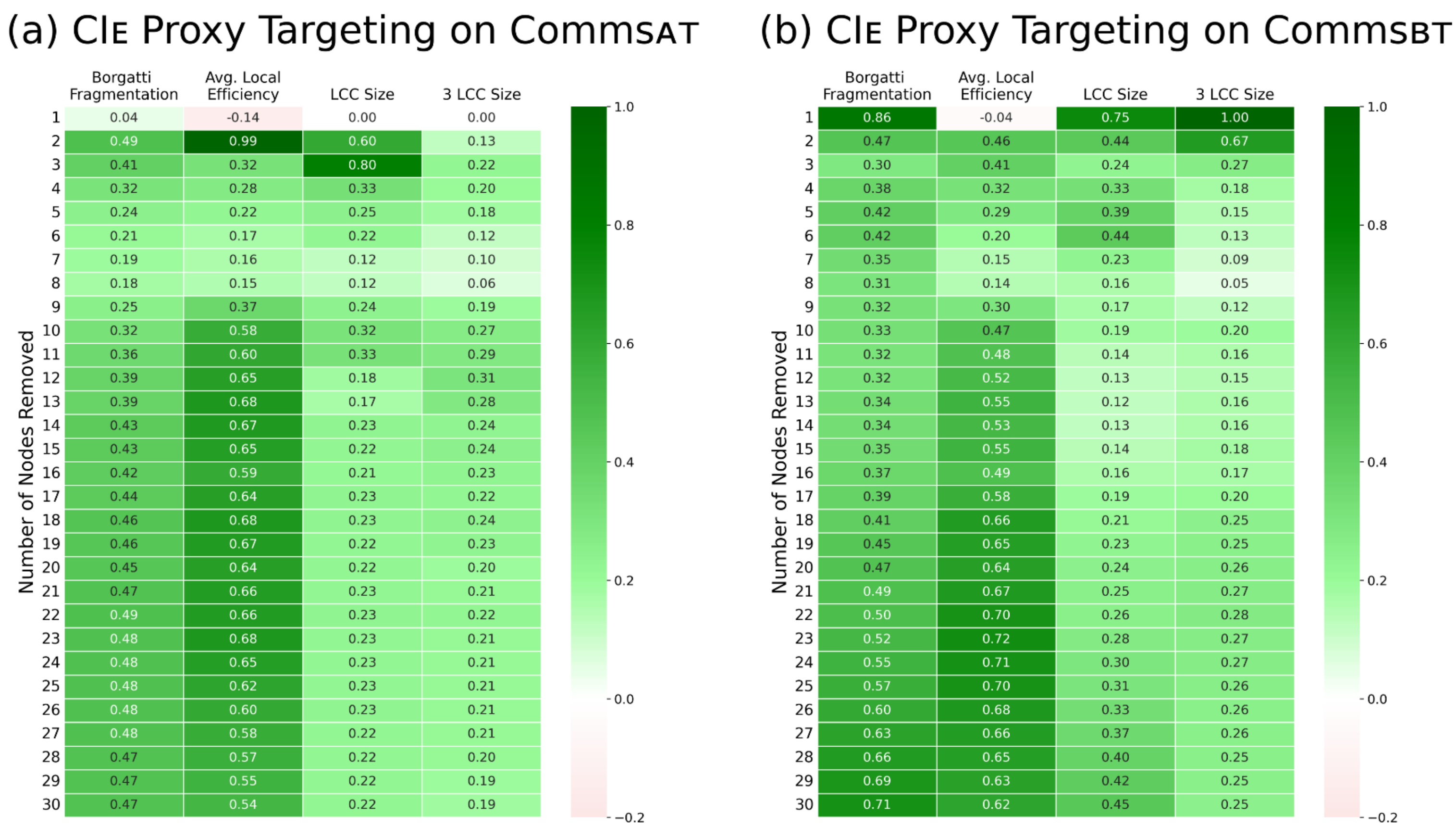


Figure 5.9. Heatmaps for the results of CI_E proxy targeting on the $Comms_{AT}$ (left) and $Comms_{BT}$ (right) dimensions.

Based on Figure 5.8, removing nodes that are prominent in the communication dimensions tends not to be as effective at disrupting the CI_E layer as native targeting for the first 10-15 nodes; for larger sets of removed nodes, however, proxy targeting based on the communication dimensions becomes more effective, especially at increasing the Borgatti fragmentation and reducing the average local efficiency of the CI_E layer that represents co-involvement in legitimate businesses. After 20 or more nodes are removed, the effect sizes for Borgatti fragmentation and average local efficiency are generally 60-90% as large as comparable effect sizes based on the native targeting disruption analyses; this drops to 40-80% when considering the share of nodes in the largest connected component and three largest connected components.

On the other hand, as presented in Figure 5.9, removing actors that are central in the CI_E dimension is less impactful on the overall structure of the communication dimensions. Similar to Figure 5.8, the first 10-15 nodes removed based on the proxy targeting are much less effective than the same number of nodes removed based on the native targeting disruption analyses. When using CI_E as the targeting layer, however, larger sets of removed nodes are less effective than when using the communication dimensions. For node removal sets of 20 or more nodes, proxy targeting effect sizes are generally 45-70% and 55-70% as effective as native targeting for Borgatti fragmentation and average local efficiency respectively, below comparable results when using the CI_E dimension as the disruption layer and the communication dimensions as targeting layers. This further drops to 20-45% and 20-30% for the shares of nodes in the largest connected component and the three largest connected components.

5.5 Discussion

The newly-proposed proxy targeting approach enabled new sociologically, criminologically, and operationally relevant research questions to be answered. The first set of analyses – investigating the usefulness of proxy targeting using the publicly-available CI_O dimension to identify nodes for cohesion and disruption analyses – aimed to test whether business register network data could help law enforcement agencies target surveillance and disruption of mafias and other organized crime groups that have infiltrated legitimate businesses. Results suggest the business register-based corporate interlock dimension acts as a good proxy for targeting the police intelligence-based corporate interlock dimension that also contains covert ties not visible in publicly-available sources, especially when targeting smaller sets of up to ten nodes. While the effectiveness of the proxy targeting decreased for larger sets of nodes, prior research suggests that the first few targeted nodes are especially important and drive the majority of the disruption (Jia et al., 2024, p. 5), a result supported by the native targeting disruption results presented in Section 5.4.2; a dynamic disruption analysis using agent-based models similarly suggests that the removal of five or ten nodes can lead to the overall dissolution of an organized crime group, with marginal additional benefit from larger removal sets (Manzi & Calderoni, 2024a, 2024b). Thus, targeting strategies should focus on identifying a limited set of key actors on which to focus a police investigation for maximum impact rather than targeting large subsets of a group. Outside of targeting a removal set of one node – where proxy targeting based on the CI_O dimension was as or more effective than native targeting for the CI_E dimension – proxy targeting generally underperformed native targeting; however, this must be balanced with the cost of acquiring the data used in each analysis. The proxy targeting is based on the CI_O dimension, the data for which comes from a publicly-available business register: this type of information is inexpensive or freely accessible, easy to collect, and does not require official authorization. On the other hand, the native targeting in this case was based on the CI_E dimension – while this dimension is richer because it also includes covert relationships and native targeting on it outperforms the proxy targeting, the electronic and physical surveillance evidence on which the network data were based is expensive and time-consuming to collect and not available to law enforcement at the outset of an investigation. The findings in Section 5.4 suggest that, while imperfect, proxy targeting using business register information can be a valuable tool for law enforcement to prioritize targets to surveil and focus on at the outset of an investigation into organized crime groups in legal markets.

Second, the analyses in this paper sought to answer the question of whether network multiplexity can provide redundancy that reinforces connectivity against targeted node removal or whether disrupting one dimension in a multiplex network leads to the destabilization of other dimensions. Prior studies suggest that multiplex ties are key for connecting a network (C. M. Smith & Papachristos, 2016) and that removing key nodes can fragment multiple dimensions of a multiplex network (Malm et al., 2010; Toledo et al., 2023). In this case, however, results suggest that multiplexity acted as a mechanism for reinforcement and resilience. While proxy targeting on the CI_O dimension was effective at disrupting the CI_E dimension, it was largely ineffective at destabilizing the communication dimensions, $Comms_{AT}$ and $Comms_{BT}$. Further, as seen in Section 5.4.4, proxy targeting using the CI_E and communication dimensions was ineffective for smaller sets of removed nodes – for example, removing a small number of nodes from the CI_E dimension that were prominent in the $Comms_{AT}$ dimension was less effective than removing the same number of nodes from that layer based on their prominence in the CI_E dimension. As the number of nodes removed increased, proxy targeting became more effective for both the CI_E and communication dimensions; however, removing larger sets of nodes prominent in the communication dimensions from the corporate interlock dimension was more effective than the reverse.

These results lead to two main conclusions. First, multiplexity can enhance resiliency by creating new dimensions that can maintain connectivity even when others fail. This is a crucial consideration for law enforcement and academic researchers when considering disruption strategies and organized crime groups. Focusing solely on one dimension could obscure that another layer of the network has been unaffected by node removal. This is also an important consideration for non-criminal organizations that are trying to enhance their ability to withstand and respond to external shocks – multiplexity can be an advantage rather than a liability. However, the resilience afforded by multiple tie types can differ by the type of shock – in the case of the results in Section 5.4.4, multiplexity reinforced connectivity for small sets of removed nodes, but became less effective at doing so when larger number of actors were taken out.

Second, different dimensions can provide different levels of multiplex resilience. The results presented in Section 5.4.4 suggest that the communication dimensions are more effective at providing backup connectivity and resiliency than the corporate interlock dimension CI_E for large numbers of removed nodes. These sorts of results are key for understanding which types of relationships are critical for enabling multiplex resilience within a network.

In addition to the substantive results, these analyses show the utility of the proxy targeting approach. Both of the research questions investigated would be difficult to answer using the standard native targeting approach – the analyses required more network-based information than what is contained within a single dimension or monoplex network. Beyond these specific questions and the context of the case in the analyses, there are further sociologically and operationally relevant topics that were previously not addressable because the standard methods could not accommodate them. The nearest extension is to understanding the resilience of non-criminal social networks. While multiplexity reinforced overall connectivity in the case of the `Ndrangheta network investigated in this analysis, non-criminal organizations like corporations or social movements may exhibit the frailty suggested by prior studies, where the removal of key nodes damages multiple dimensions at once (Boccaletti et al., 2014). To develop resilient organizational structures, it is necessary to know whether such vulnerabilities exist and how different dimensions of connections can interact to bolster them. This approach could also be applied to non-social networks that exhibit multiplexity, such as urban mobility and transportation networks; a better understanding of how the inoperability of a key transportation interchange that interacts with multiple mobility channels like trains and buses impacts each of those channels can allow for the engineering of more resilient systems.

5.6 Conclusion

Prior network disruption analyses have helped improve our understanding of what makes networks resilient and how law enforcement can target organized crime groups to destabilize them and their criminal operations. The approach used to identify nodes for removal, however, constrained the ability of the analyses to account for multiplexity and the type of information available to law enforcement agencies, limiting their sociological, criminological, and operational utility. This paper has proposed a new approach – proxy targeting – that separates the targeting and disruption calculation functions. In doing so, this approach better accounts for the multiplex and multidimensional nature of many organized crime group networks and allows for the inclusion of richer and more easily accessible network data for targeting. To demonstrate the utility of this approach, the analyses in this paper aimed to answer two main research questions.

First, when analyzing a mafia network that has infiltrated legitimate businesses, can business register network data be useful for targeting actors for surveillance and arrest? Results show that targeting based on these publicly-available network data, while not as effective as using police intelligence-based corporate interlock and communication dimensions to target themselves, is approximately 60% as effective for targeting surveillance and 55-100% as effective when targeting the intelligence-based corporate interlock dimension for disruption using a small removal set size, even outperforming the native targeting in a small number of cases; as an easy-to-collect and inexpensive data source, business register network data provide an opportunity for law enforcement to prioritize resources and guide surveillance in the early stages of an investigation. On the other hand, proxy targeting using business register data was not effective at disrupting the communication dimensions.

Second, an analysis was conducted to determine whether multiplexity bolsters the resiliency of organizations or increases the risk that targeting key nodes could destabilize multiple dimensions simultaneously. Using the police intelligence-based corporate interlock and communication dimensions, a proxy targeting network disruption analysis revealed that removing a small number of nodes that are central in one dimension from another dimension was generally not effective at disrupting that second dimension; however, the communication dimensions provide more reinforcement to the connectivity of the corporate interlock dimension than vice versa. Further, removing larger sets of nodes was more effective at destabilizing the different dimensions in the multiplex network.

These results can help inform how law enforcement agencies identify actors to investigate, how academic researchers can better understand the resilience and vulnerability of criminal networks, and how non-criminal organizations and systems can be designed to withstand external shocks. Further, the proxy targeting approach opens new avenues for future research. First, the analyses in this paper only included structural targeting metrics and a limited set of performance metrics and did not allow for post-disruption adaptation. Future analyses could build on this approach using dynamic simulations like agent-based modelling (Manzi & Calderoni, 2024a, 2024b) and empirical post-disruption examples (Berlusconi, 2022; Diviák et al., 2022), as well as more complex targeting and performance criteria, especially those that explicitly account for multiplex centrality and connectivity. Second, a wider range of organized crime group networks can be analyzed using this approach; for example, previously-analyzed data on multiplex criminal networks provided by Ficara et al. (2021) and Toledo et al. (2023) could be used to further test the efficacy of the proxy targeting approach. In particular, understanding how this new approach functions under constraints like incomplete and noisy data would validate its use in both operational and academic settings, similar to the approaches of Yeung et al. (2025) for disruption analyses and Berlusconi (2013) for organized crime group network data more generally. Regarding new research questions, proxy targeting of a mafia network using a network representation of the group's formal hierarchy could provide a more nuanced understanding of whether social networks in mafias converge with or diverge from formal organizational structures and could add additional information to the role-based targeting strategies proposed previously in the literature. Finally, the proxy targeting approach could be extended to analyze non-criminal networks like those in corporations, governments, and social movements to better understand whether and how these structures can use multiplexity to avoid destabilization.

6

Conclusion

6.1 Summary and Implications

In this thesis, I have presented analyses of the structures of three mafias using novel data sources. In the first paper, we used public media reporting of violent yakuza-on-yakuza attacks in Japan to understand the pattern of violence amongst yakuza groups and uncover what these patterns suggest for the structure of yakuza syndicates. While syndicates have formal authority over their constituent groups based on the yakuza's complex formal hierarchy and multilevel divisional organizational structure, prior ethnographic accounts have disagreed about the extent to which this control exists in practice. Using newly-developed analytical techniques and a multilevel exponential random graph model, we found that unlike street gangs in North America (Decker & Curry, 2002; Randle & Bichler, 2017), yakuza groups identify themselves with their syndicate: groups that are part of the same syndicate tend not to attack one another and retaliate on each other's behalf. Revealingly, this retaliation is typically not against the original attacker directly, but against another group in the attacker's syndicate – this suggests that yakuza groups not only identify with their own syndicate, but also associate other groups with their respective syndicates such that an attack by a group can be properly avenged by attacking another group in its syndicate. On the other hand, while yakuza syndicates form alliances and sometimes create multi-syndicate alliance structures, these positive relationships reduce the likelihood of attack between allied groups but fail to create the same sense of shared identity, as evident from the lack of reciprocal attacks on behalf of allies. Finally, yakuza groups are more likely to attack other groups in the same city and prefecture. While syndicates function as cohesive organizations with a shared identity, decisions to attack made by local yakuza groups are still driven in part by local market factors.

These findings contribute to the literature of mafia and organized crime group conflict, revealing materially different patterns and dynamics than those seen amongst street gangs. More broadly, this paper makes two key contributions for understanding the structure of mafias, other organized crime groups, and non-criminal organizations. First, it stresses the importance of multilevel considerations. Mafias commonly share a multilevel divisional structure, with the main operational units – often called 'families' – belonging to higher-level groups, which are then part of an overarching organization through one or more intermediary levels (Catino, 2019; Varese, 2017). Other organized crime groups and legal organizations like corporations and universities often share this structure, where members are part of both a lower-level group and the overall organization, like an academic researcher at a sociology department in the social sciences division of a university. While some studies on organized crime group structure have sought to explicitly model these considerations (e.g., Coutinho et al., 2020), many analyses of mafia structure either fail to include this dimension or use group membership as a node attribute without incorporating information on the relationships between groups. The case of yakuza conflict, however, shows that failing to properly account for multilevel divisional considerations can yield misleading results. When only viewing the conflict network between local yakuza groups, there appeared to be no structure or pattern to the attacks; the mechanisms behind conflict dynamics only become clear when the affiliation of groups to syndicates and the alliances between syndicates are considered.

Second, these findings suggest that yakuza syndicates exhibit a convergence between formal and informal structures. Formal structure represents the hierarchy, distribution of authority, and relationships between bureaucratic positions within an organization, while informal structure constitutes the non-formalized aspects like power dynamics, social expectations, and interpersonal relationships that arise as criminal activities are carried out (Catino, 2019, pp. 135, 151; see also Barnard, 1938). These organizational maps can converge – with formal authority correlating with social power and interpersonal relationships aligning with the connections between the roles those people inhabit – or they can diverge. Prior studies of mafias have suggested a divergence between the formal and informal structures in the Sicilian Cosa Nostra (Musotto, 2022), the Calabrian `Ndrangheta (Calderoni, 2012), and the Italian-American Mafia (Krajewski et al., 2022). The results in the first paper of this thesis, on the other hand, provide evidence of a convergence between the formal and informal, with the formal ties of co-affiliation to a syndicate leading to groups retaliating on each other's behalf and syndicate alliances reducing conflict rates among their local groups. While some do experience divergence between their formal and informal structure, considering a wider range of mafias reveals that not all do.

The second paper explored a similar topic within the *vory*. The *vory* exist in a formally flat association, but prior ethnographic studies suggest that the fraternity exhibits factionalization and power hierarchies, although accounts differ in their proposed explanations for the reported structure. This paper aimed to use a set of *progony* written and signed by *vory* to explore the pattern of relationships between these mafia bosses and determine the drivers and mechanisms behind tie formation. Results based on the *progony* co-signing networks suggest that the *vory* are polycentric and factionalized, with clusters built around multiple sets of key mafiosi rather than a clear core and periphery. While prior accounts have portrayed the fraternity as disconnected, we found evidence of cohesion, with interconnections between certain topological communities and brokers bridging between different factions to hold the network together. Moving beyond descriptions of the structure of the *vory*, we investigated the drivers of co-signing. Rather than *vory* co-signing with other bosses of the same ethnicity or of similar age, we found that co-signing behavior was most associated with prior crowning ties – a *vor* was more likely to sign a *progon* with the crowner that sponsored them to receive the title. While shared ethnicity did have a positive and statistically significant impact on co-signing likelihood, the magnitude of this effect was only one-ninth as strong as the effect for crowning ties. Further, results suggested that ethnicity and age are not correlated with crowning ties – *vory* do not preferentially crown mafiosi of the same ethnicity or age.

The findings in the second paper represent the first direct analysis of the relationships between mafia bosses within a higher-level coordination body. Prior study of this level of mafias has been difficult due to a lack of data. Higher-level coordination bodies and boss-to-boss interactions in most mafias are infrequent and difficult for law enforcement to infiltrate or surveil, meaning the police and judicial data sources commonly used in academic research cannot adequately capture the network structure of this organizational level. Instead, we used publicly-available information and benefited from the fact that the *vory* publish signed documents in a novel approach to studying mafia structure.

Further, this paper contributes to the literature by illustrating a new way in which the formal and informal structures of a mafia can diverge. Unlike the first paper that found a convergence between the formal hierarchy of the yakuza and the informal structures evident from patterns of conflict, this paper finds divergence for the *vory* – the official egalitarianism of the association masks the underlying interpersonal power imbalances, pecking order, and factionalization. This is similar to how prior research on other mafias has found a divergence between formal and informal organizational structures in the lower levels of other mafias. A key distinction, however, is how the structures diverge. In prior studies that found divergence between the formal and informal in the Sicilian Cosa Nostra and the Calabrian `Ndrangheta, the divergence occurs because the highest-ranking mafiosi – the bosses and other high-ranking members – were not the most central actors in operational networks relating to criminal activities; instead, mid-ranking members – those that sit between the bosses and the lower-ranking foot soldiers of the organization – were the best connected members of the networks (Calderoni, 2012; Musotto, 2022). This could arise from higher-ranking members both preferring and being able to limit their involvement and centrality to avoid detection and arrest by law enforcement or from mid-level mafiosi being tasked more with operational duties, while bosses instead focus on the management of the overall group; both explanations would suggest that the divergence is driven by differences in the formal authority and informal centrality of a given role in the mafia hierarchy. The *vory*, on the other hand, all share the same title and the same formal role and, as such, are on equal footing when interacting with each other within the fraternity. Even in this case, certain *vory* have emerged as poles of power around which others cluster, while other *vory* have taken brokering positions between different communities in the network. Rather than role-based divergence, another set of mechanisms must be at work.

Future research is needed to determine what led certain *vory* to the special network positions they inhabit.

The third and fourth papers in this thesis turned from the structure of higher-level interactions within mafias to focus on how we can investigate how mafias are organized as they infiltrate legal markets. The third paper sought to solve data limitation issues by evaluating and testing a novel data source: publicly-available business registers. While these sources have been used when studying corporate interlock among legitimate companies, they have been underutilized when investigating organized crime groups like mafias in legal markets. One potential driver of this underutilization is a worry about reliability and accuracy: as publicly-available and government-maintained sources, business registers may miss critical information about hidden criminal activities or relationships. I sought to determine the usefulness of this new data source by comparing it against a more commonly-used one, an Italian pre-trial detention notice. Overall, I found that business register-based corporate interlock networks can be useful for studying mafias and organized crime groups in legal markets, but that they must be applied carefully to ensure validity.

Despite lacking direct evidence of criminal interactions – like those that might be visible in closed sources like an indictment – the business register data both proved useful as a supplement to law enforcement data and contained further detailed data that would allow it to function as a standalone data source on ties between actors. The pre-trial detention order contained erroneous information on the official ownership of and affiliation to companies for 13% of the businesses mentioned in the order; in this case and others like it, the business register information should be used to certify and correct the information in the law enforcement data prior to analysis. Further, the business register data contain additional information not available in the indictment. For example, the business register contains all of the actors involved with the infiltrated companies rather than the subset that the police and prosecutors deemed relevant to the legal case against the `Ndrangheta group; the publicly-available data also contained more information on the sector and corporate governance of the relevant companies, the relationship between businesses, and on how the network changed both over time and after disruption by law enforcement. This additional detail allows researchers to answer a wider array of questions than those addressable using the more limited closed-source data, such as whether mafiosi are industry-agnostic or instead specialize in certain sectors, how mafias interact with non-criminal business owners and collaborate with enablers like accountants and lawyers, and whether mafia structures in the illegal market are replicated in the ownership structure of legitimate businesses.

Both academic and operationally-focused researchers, however, are also interested in hidden criminal structures. Closed sources like police and judicial data have traditionally been used to investigate these structures given their access to non-public evidence from electronic and physical surveillance. These sources, however, are not always accessible. When this is the case, the results from this paper suggest that publicly-available business register data can approximate the view provided by closed-source data. However, this is only true for certain levels of analysis and network types. The business register-based corporate interlock network failed to identify central actors or replicate whole-network structural characteristics in the police data-based corporate interlock and communication networks. On the other hand, the business register-based and police data-based corporate interlock networks shared similar local neighborhood-level structures. Beyond being useful as a supplemental data source and as a means of investigating specific research questions that do not hinge on knowledge of hidden criminal relationships, this suggests that business register-based network data can be used as a proxy when investigating tie formation mechanisms in criminal corporate interlock networks using models like exponential random graph models and stochastic actor-oriented models. This paper contributes to the study of mafias, organized crime groups, and legal market infiltration by evaluating a new data source and determining analyses and topics for which it is well-suited, as well as demonstrating areas for which it should not be used. Further, it highlights the importance of testing novel data sources before use and provides a possible template for evaluating other data sources to diversify the range of organized crime groups, contexts, and questions that can be investigated.

Utilizing the network data tested in the third chapter, the fourth paper used a newly-proposed proxy targeting technique for network disruption analyses to investigate multiplex resilience in a mafia network. The two aims of the study were to (1) evaluate whether publicly-available business register network data could be used to target surveillance and disruption of criminal networks in the legal economy and (2) investigate the impact of multiplexity on network resilience against disruption. This paper contributes to the literature on mafias in legal markets and network resilience by proposing a new methodology that enables new empirical and sociological findings.

Prior organized crime network disruption analyses used a native targeting approach, where nodes were chosen for removal based on their position in the underlying network of interest; however, this requires complete knowledge of the network based on electronic and physical surveillance, meaning it cannot be used operationally by law enforcement agencies to design targeting strategies. Using the proxy targeting approach, I investigated whether targeting nodes that were central in the business register-based corporate interlock dimension was effective at selecting nodes for surveillance and removal from police data-based corporate interlock and communication dimensions in a multiplex network. For identifying nodes that are well-connected in the dimensions of interest and that could provide a clear view of the network when surveilled, proxy targeting was generally 55-65% as effective as the optimal targeting based on the underlying dimensions themselves. Further, removing nodes that were central in the publicly-available corporate interlock dimension was 55-80% as effective at reducing whole-network connectivity as native targeting in the police data-based corporate interlock dimension for up to five nodes removed; however, proxy targeting based on the publicly-available corporate interlock dimension was largely ineffective at disrupting the two communication dimensions. In both the cohesion and disruption analysis, proxy targeting based on publicly-available data was less effective than native targeting on the police data-based dimensions. The police data, however, is based on evidence collected through long and expensive electronic and physical surveillance operations. Network data from business registers, on the other hand, are inexpensive or free, easy to access and collect, and do not require authorization for wiretaps or physical surveillance. Business register data can offer a means of identifying and prioritizing actors in the early stages of a law enforcement investigation into infiltration of legal businesses by mafias and other organized crime groups.

Multiplexity – when actors are bound by multiple types of relationships in different social contexts – has been understudied regarding network disruption and resilience, in part due to a lack of suitable data and analytical techniques. When investigating whether the removal of nodes that are central in one police data-based dimension from another dimension is effective at disrupting that second dimension, findings suggest that removing a small set of nodes did not destabilize the network. For central nodes in the police data-based corporate interlock dimension, removal from the communication dimensions failed to materially increase their fragmentation. Removing larger sets of nodes based on proxy targeting, however, became more effective, although not as effective as native targeting. In sum, these results go against prior conclusions in the literature that multiplexity can create key points of failure across multiple dimensions (Malm et al., 2010; Toledo et al., 2023). Instead, in this case multiple tie types reinforced connectivity – when key nodes were removed from one dimension, the other dimensions were largely unaffected. This was especially true of the communication dimensions, which were more resilient against removal of nodes that were central in the corporate interlock dimension than vice versa. These results suggest that disruption analyses of mafias and other organized crime groups, as multiplex organizations, must consider the different dimensions in each network instead of focusing on a single social relationship. Further, it is not the case that multiplexity always reinforces or weakens resilience: beyond varying across cases and groups, it can also differ by the dimension considered. These findings stress the importance of considering multiplexity generally and determining the differential impact of different dimensions rather than speaking about multiplexity in broad terms.

6.2 Avenues for Future Research

The papers in this thesis suggest several avenues for future research. The first paper finds differences in the pattern of violent conflict between yakuza groups compared to prior trends in the literature on street gangs due to the differences in their internal organizational structure. While other mafias are more similar to the yakuza than to street gangs, each organization has its own set of internal processes and dispute resolution infrastructure. Further study on conflict within and between mafias and other well-structured organized crime groups is needed to understand the extent to which the findings in this paper extend to other similar organizations and the impact of different institutional differences. For example, do mafias with weaker internal controls than the yakuza see more complex conflict patterns?

More broadly, the results in this chapter highlight the importance of considering multilevel divisional structure within mafias and other organized crime groups. Many studies on the organization of mafias fail to account for this structural characteristic at all. Others incorporate mafia 'family' membership as a nodal attribute assigned to individual mafiosi (Calderoni et al., 2017; DellaPosta, 2017, 2023); while this approach brings in aspects of the multilevel divisional structure of mafias, there are more complex dependencies above the level of family membership. In a multilevel divisional structure like that observed in mafias, there are both within-level and between-level connections. Subgroups at the same organizational level can interact in both positive and negative ways, like yakuza syndicates forming alliances and local yakuza groups attacking each other; this level has seen some investigation in recent years (Catino et al., 2022). These organizational units, while part of an overarching superstructure, have their own priorities, capabilities, resources, and relationships; while united towards a common goal, they can also compete and cooperate with each other to achieve their own aims. There are also relationships between groups at different levels, like a local yakuza group receiving orders from its syndicate. Different organizations set out different procedures for interactions between levels. Further, these two types of interaction can overlap, as in the case where a local group escalates a dispute to its syndicate, which then corresponds with its ally on the other side of the dispute. More research is needed to understand these complex within- and between-level interactions within mafias.

For example, one area for additional study is the convergence or divergence of the formal and informal aspects of multilevel divisional structures. While the formal structure of an organization can set out how subgroups at the same level should interact and how groups at different levels relate to one another, this can converge with or diverge from how relationships actually form and how activities are carried out. The results in the first paper suggest that the yakuza sees convergence between the formal and informal multidivisional structures, where the prescribed organizational setup matches what is observed in social interactions between its constituent members and groups. Other mafias or criminal groups, on the other hand, may exhibit signs of divergence, where the formal hierarchy of the different organizational levels does not impact the real-world interactions between the components of the multilevel system.

The second paper used *progony* co-signing decisions to understand the structure of the *vory* association and the drivers of interactions between mafia bosses. While instructive, these co-signing decisions are just one of the relationships between different *vory*. To better understand the structure of the fraternity, a next step would be to gather additional network datasets to enrich our understanding of the relationships between the *vory*, with a focus on multiplexity – for example, other tie types could include co-offending and co-arrest ties, mutual presence at publicly-visible social events like funerals and birthday parties, and negative relationships relating to violent conflict or de-crowning attempts. These additional data could build on the view of the *vory* presented in this paper.

Beyond the 'Russian Mafia', similar investigations using non-standard data could also prove fruitful for investigating the structure of higher-level coordination bodies in other mafias. While commonly-used police and judicial data sources have historically offered little insight into the upper echelons of mafias, this paper shows that novel sources can enable research into this organizational level. Beyond advancing our understanding of other mafias, this would also allow comparison between different mafias. This would allow for further investigation of the relationship between formal and informal organizational structures: while the formally-egalitarian *vory* exhibit signs of factionalization and vertical power relationships, other mafias with different higher-level organizational characteristics may instead see convergence between how bosses ought to relate to one another and how they actually do. Further, comparison between mafias would allow us to better understand how the structure of higher-level coordination bodies and the patterns of interactions among mafia bosses evolve over time. In particular, as other mafias have faced increased law enforcement pressure and subsequently lost their higher-level coordination bodies or seen their use curtailed, they may become more similar to the *vory*. By studying the differences and changes within this level of mafias across a range of organizations, we can better understand what factors and constraints impact how mafia bosses collaborate and compete.

Comparison between the *vory* and non-criminal organizations could allow us to test theories of faction formation and brokerage in a novel context and better understand the mechanisms behind these social processes. The *vory* are not the only group to see factions and brokers; Burt (2004), for example, examines key brokers between divisions within the network of mid-level supply chain managers in an electronics business. One topic for further research is understanding how factions and brokers between them arise in different contexts, such as legal and illegal settings. Further, the 'factions' within Burt (2004)'s work were institutionally-defined business units, while the *vory* have no such formal distinctions in their association – comparison between these cases and those like them would allow for a richer understanding of why and how brokers emerge in the first place.

The findings in the third paper enable multiple extensions. First, other business registers and case studies should be evaluated in a similar manner to confirm whether the results generalize to other contexts; for cases where business register data perform better or worse than in the case presented in this paper, characteristics of the business register and the closed-source data could be compared to understand in which contexts we can expect business register corporate interlock network data to function well as a standalone source or as a proxy for closed-source data. Second, business registers like Il Registro Imprese can be used in substantive analyses of the structure of mafias and other organized crime groups as they infiltrate legitimate businesses. Using this new data source can help researchers understand industry specialization and collaboration with entrepreneurs and facilitators, as well as compare the structures of mafias in the legal and illegal markets. The results from this chapter also lay out the limits of this data source and suggest analyses and topics for which it ought not to be used. Third, other open-source and publicly-available data sources can be evaluated in a similar way. A key driver of our limited understanding of mafia structure has been the small set of data sources used. Expanding the sorts of sources can help broaden our view; however, it is important to confirm that the data source accurately and reliably represents the social phenomena of interest. One strategy would be to examine the data sources and methods used in other fields like economics, political science, history, and organizational theory to study non-criminal organizations; importing insights from other areas can provide new ways of investigating mafias and other criminal groups. Finally, one limitation of the third paper is that the collection of publicly-available information was based in part on a list of companies identified as relevant from the pre-trial detention order. To better understand whether this data source would be useful in situations where no official, closed-source data is available at all, a fully open-source data collection cycle should be evaluated. For example, rather than using a list of companies identified by law enforcement, suspicious and infiltrated businesses could be identified in public media sources or by using graph mining techniques on the corporate interlock network of the entire business register.

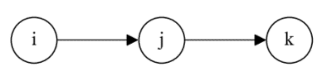
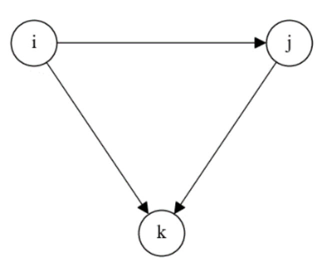
The fourth paper on multiplex resilience and publicly-available data targeting for network disruption opens two principal avenues for future research. First, a number of modelling choices were made in the disruption analyses. A limited set of targeting criteria and performance metrics were chosen which future analyses could add to or change; in particular, targeting criteria that consider non-structural node characteristics and incorporate information on multiplex positioning could extend the analyses beyond the largely monoplex structural measures used in this paper. Further, the disruption analyses in this chapter are static: performance metrics are calculated on the disruption dimension after nodes are removed without allowing for the network to adapt. A dynamic approach to disruption that considers network responses to destabilization can be incorporated into the proxy targeting approach so that results could better align with empirical outcomes. Second, the concept of multiplex resilience has been understudied in the existing literature. In contrast to prior claims that multiplexity could weaken networks by creating key nodes who are instrumental to the connectivity of multiple dimensions, the results in this paper suggest that the corporate interlock and communication dimensions of the `Ndrangheta network are both resilient to the removal of nodes that are central in the other dimension, to varying degrees. Further analyses that explicitly model multidimensional and multiplex disruption – using the proxy targeting protocol or other approaches – in both criminal and non-criminal networks can lead to insights into under what conditions multiplexity harms or builds resilience.

Appendices

A

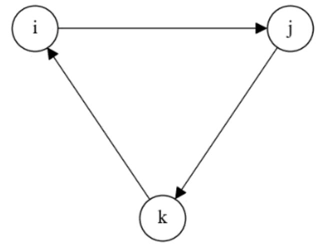
Visualizations of status-oriented conflict network configurations

Table A.1 Network Configurations

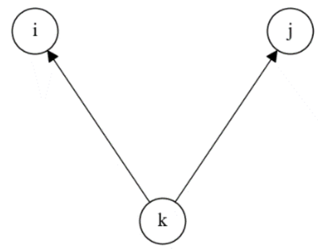
	Configuration
Retaliation	
Two-Path	
Transitive Triad	

A. Visualizations of status-oriented conflict network configurations

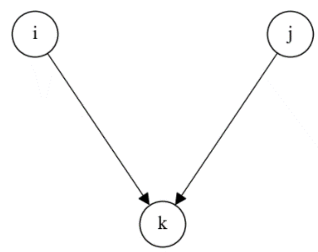
Generalized Exchange



Out-Star



In-Star



B

Yakuza conflict network triad census

Table B.1 Triad Census

Configuration	Out-star	In-Star	Two-Path	Transitive Triad	Generalized Exchange
Triad Census Label	021D	021U	021C	030T	030C
Count	1	12	15	0	0

C

Robustness of *progony* co-signing network structural characteristics to projection method

The count overlap projection method was used to convert the 2-mode *vor*-to-*progon* network into the weighed 1-mode *vor*-to-*vor* network, where the weight of an edge represents the number of documents that a pair of actors both signed. This projection method was chosen because it allows for more intuitive interpretation of edge weights than other projection methods. To show that the results are robust to projection method and not just the result of the choice to use count overlap to project the 2-mode network, the results of the MRQAP models and normalized rich-club coefficient distributions for the weighted 1-mode network generated by using the Newman projection method on the 2-mode network are presented in Table C.1 and Figure C.1 respectively. Proposed in Newman (2001), the Newman projection method adjusts the weight of ties between nodes in a 1-mode projected network based on the number of other nodes that were also connected to a given node of the other mode. Using the example of academic paper co-authorship networks where ties exist between authors and papers, the count overlap projection method would create an edge between two authors with a weight equal to the number of papers they co-authored together; using the Newman projection method, on the other hand, if two authors co-authored a paper together that had *n* total authors, that paper would contribute a value of $1/(n - 1)$ to the weight of the edge between the two authors in question, with the full weight of the edge between the two authors equal to the sum of that value across all papers they co-authored together (Newman, 2001). On this approach, many actors collaborating on a single event – like many *vory* signing one *progon* – decreases the value of the collaborative tie between any two of the actors.

The normalized rich-club coefficient distributions in Figure C.1 reveal similar trends to the results shown in Figure 3.5a-b in Section 3.4.2 – both plots show evidence of strong weighted rich-club ordering, suggesting that when well-connected actors co-sign together, they form stronger ties than expected at random. The MRQAP results in Table C.1 similarly support the results in Table 8 in Section 3.4.3. While using the Newman projection method results in the Age Difference variable becoming statistically significant in all models, the effect sizes of both age similarity and shared ethnicity are marginal, suggesting that while there may be some effect, neither variable is a material driver of co-signing behavior. Crowning relationships, on the other hand, remain the most impactful variable in the model. These results suggest that our findings in Section 3.4 are robust to the choice of projection method used to generate the weighted 1-mode network.

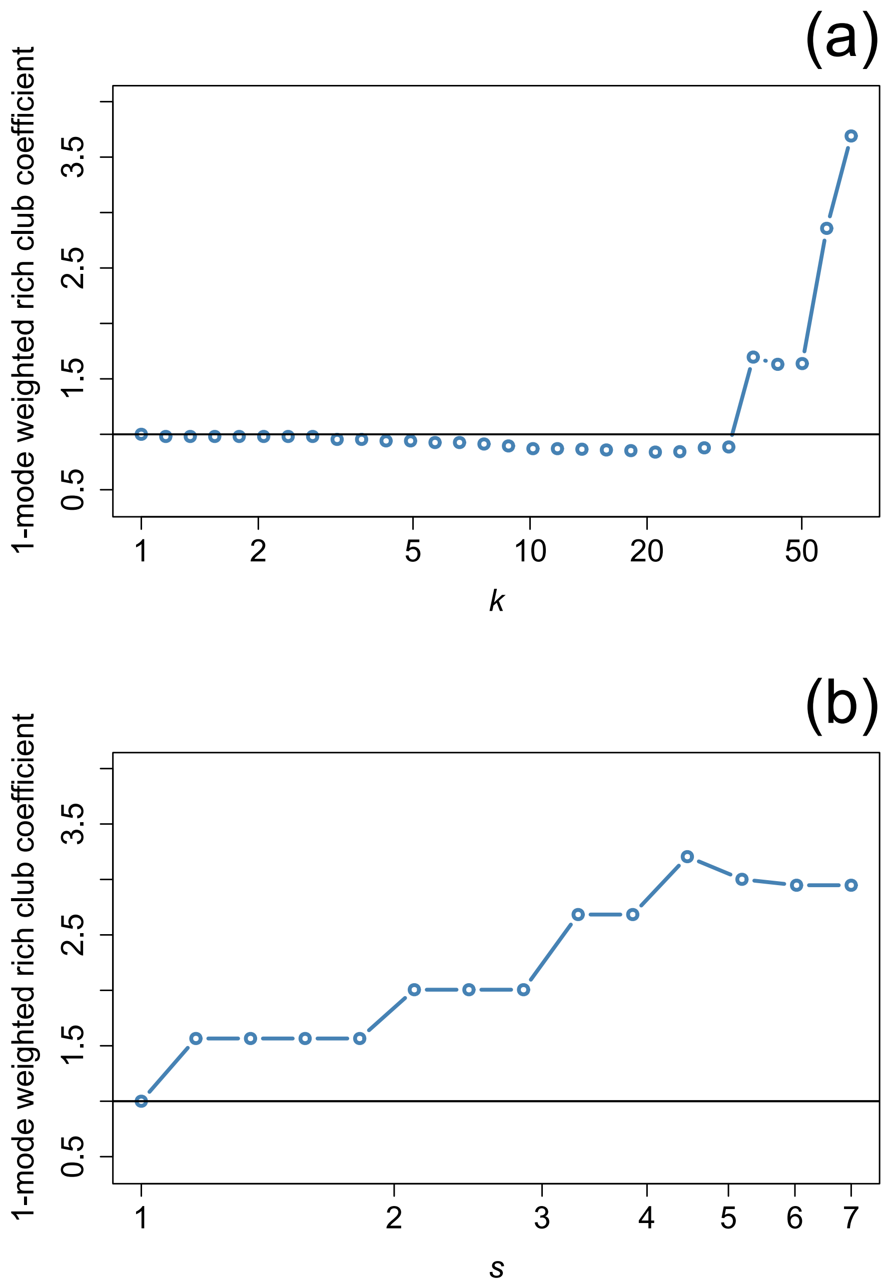


Figure C.1. Rich-club coefficient distributions for the weighted 1-mode *vor*-to-*vor* network projected using the Newman method where prominence is determined by (a) a vor's degree k and (b) a vor's strength s. Horizontal axes are presented on a logged scale.

Table C.1 MRQAP Results – Newman Projection

	Model 1		**Model 2**		**Model 3**	
Variables	Coefficient	SE	Coefficient	SE	Coefficient	SE
Intercept	0.0104***	0.0000	0.0103***	0.0000	0.0096***	0.0000
Shared Ethnicity	0.0021**	0.0010	--	--	0.0020**	0.0010
Age Difference	-0.0001*	0.0001	-0.0002**	0.0001	-0.0002**	0.0001
Crowning Relationship	--	--	0.0264***	0.0018	0.0264***	0.0018
Model Fit						
R^2	0.001***		0.020***		0.020***	
R^2 (adj.)	0.001***		0.020***		0.020***	

***p<0.01, **p<0.05, *p<0.1

D

Unweighted and weighted central node comparisons between 1-mode `Ndrangheta networks – Top 25

Table D.1 Unweighted Central Nodes Comparison Between 1-Mode `Ndrangheta Networks – Top 25

	Degree Centrality		**Betweenness Centrality**		**Closeness Centrality**		**Eigenvector Centrality**	
	In OCI Nodelist	*Overlaps OCI Top 25*	*In OCI Nodelist*	*Overlaps OCI Top 25*	*In OCI Nodelist*	*Overlaps OCI Top 25*	*In OCI Nodelist*	*Overlaps OCI Top 25*
ECI Full	80%	4%	76%	32%	72%	16%	68%	0%
ECI Core	76%	4%	72%	32%	68%	16%	60%	0%
Comms. (All Ties)	28%	4%	32%	8%	36%	4%	36%	0%
Comms. (Business Ties)	40%	4%	40%	12%	44%	4%	40%	0%

D. Top 25 central node comparisons between 1-mode `Ndrangheta networks

Table D.2 Weighted Central Nodes Comparison Between 1-Mode `Ndrangheta Networks – Top 25

	Weighted Degree Centrality		**Weighted Eigenvector Centrality**	
	In OCI Nodelist	*Overlaps OCI Top 25*	*In OCI Nodelist*	*Overlaps OCI Top 25*
ECI (Full, CO)	84%	4%	92%	0%
ECI (Full, NW)	72%	48%	100%	8%
ECI (Core, CO)	80%	4%	68%	0%
ECI (Core, NW)	60%	48%	76%	8%

Sources

Interviews:

S1: Yakuza Group Boss (2023, May) [Personal Interview, conducted by Dr. Martina Baradel].
S2: Chief of Prefectural Police (2023, July) [Personal Interview, conducted by Dr. Martina Baradel].
S3: Former Legal Advisor to Major Yakuza Syndicate (2023, April) [Personal Interview, conducted by Dr. Martina Baradel].
S4: Yakuza Journalist (2023, April) [Personal Interview, conducted by Dr. Martina Baradel].
C1: Formal Penal Colony Convict Who Spent 2 Years in Prison (2022, December) [Personal Interview, conducted by Dr. Elena Racheva].

Yakuza Magazines:

Jitsuwa Jidai. Sanwa Shuppan.
Editions: JJ 04/2007, JJ 07/2007, JJ 01/2008, JJ 11/2014, JJ 02/2015, JJ 11/2015, JJ 12/2015, JJ 04/2016, JJ 10/2016

References

Adelstein, J. (2009, October 30). The Strange World of Yakuza Fan Magazines. *Publishing Perspectives*. https://publishingperspectives.com/2009/10/the-strange-world-of-yakuza-fan-magazines/

Agreste, S., Catanese, S., De Meo, P., Ferrara, E., & Fiumara, G. (2016). Network structure and resilience of Mafia syndicates. *Information Sciences*, *351*, 30–47. https://doi.org/10.1016/j.ins.2016.02.027

Airola, A., & Bouchard, M. (2020). The Social Network Consequences of a Gang Murder Blowout. *Social Sciences*, *9*(11), 204. https://doi.org/10.3390/socsci9110204

Aziani, A., Ferwerda, J., & Riccardi, M. (2022). Who are our owners? Exploring the ownership links of businesses to identify illicit financial flows. *European Journal of Criminology*, *19*(6), 1542–1573. https://doi.org/10.1177/1477370820980368

Baika, L., & Campana, P. (2020). Centrality, Mobility, and Specialization: A Study of Drug Markets in a Non-metropolitan Area in the United Kingdom. *Journal of Drug Issues*, *50*(2), 107–126. https://doi.org/10.1177/0022042619891962

Baradel, M. (2021). Yakuza Grey: The Shrinking of the Il/legal Nexus and its Repercussions on Japanese Organised Crime. *Global Crime*, *22*(1), 74–91. https://doi.org/10.1080/17440572.2020.1813114

Barnard, C. I. (1938). *The Functions of the Executive*. Harvard University Press.

Barone, G., & Narciso, G. (2015). Organized crime and business subsidies: Where does the money go? *Journal of Urban Economics*, *86*, 98–110. https://doi.org/10.1016/j.jue.2015.01.002

Battiston, F., Nicosia, V., & Latora, V. (2017). The new challenges of multiplex networks: Measures and models. *The European Physical Journal Special Topics*, *226*(3), 401–416. https://doi.org/10.1140/epjst/e2016-60274-8

Battiston, S., & Catanzaro, M. (2004). Statistical properties of corporate board and director networks. *The European Physical Journal B - Condensed Matter*, *38*(2), 345–352. https://doi.org/10.1140/epjb/e2004-00127-8

Benischke, M. H., Bhaskarabhatla, A., & Singh, R. (2024). The Effect of Incoming Board Interlocks With Public Firms on Private Firms' Survival: Large-Scale Evidence From

India. *Journal of Management*, *50*(5), 1939–1971. https://doi.org/10.1177/01492063231164971

Berlingerio, M., Koutra, D., Eliassi-Rad, T., & Faloutsos, C. (2012). *NetSimile: A Scalable Approach to Size-Independent Network Similarity* (arXiv:1209.2684). arXiv. http://arxiv.org/abs/1209.2684

Berlingerio, M., Koutra, D., Eliassi-Rad, T., & Faloutsos, C. (2013). Network Similarity via Multiple Social Theories. *ASONAM '13: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 1439–1440.

Berlusconi, G. (2013). Do all the pieces matter? Assessing the reliability of law enforcement data sources for the network analysis of wire taps. *Global Crime*, *14*(1), 61–81. https://doi.org/10.1080/17440572.2012.746940

Berlusconi, G. (2016). Organised criminals and the legal economy. In E. U. Savona, M. Riccardi, & G. Berlusconi (Eds.), *Organised Crime in European Businesses* (pp. 4–15). Routledge. https://doi-org.ezproxy-prd.bodleian.ox.ac.uk/10.4324/9781315640617

Berlusconi, G. (2022). Come at the king, you best not miss: Criminal network adaptation after law enforcement targeting of key players. *Global Crime*, *23*(1), 44–64. https://doi.org/10.1080/17440572.2021.2012460

Bichler, G., & Norris, A. (2024). Crossing lines: Structural advantages of inter-racial criminal street gang violence. *Trends in Organized Crime*, *27*, 235–262. https://doi.org/10.1007/s12117-022-09468-9

Bichler, G., Norris, A., Dmello, J. R., & Randle, J. (2019). The Impact of Civil Gang Injunctions on Networked Violence Between the Bloods and the Crips. *Crime & Delinquency*, *65*(7), 875–915. https://doi.org/10.1177/0011128717739607

Bichler, G., Norris, A., & Ibarra, C. (2020). Evolving Patterns of Aggression: Investigating the Structure of Gang Violence during the Era of Civil Gang Injunctions. *Social Sciences*, *9*(11), 203. https://doi.org/10.3390/socsci9110203

Bichler, G., Norris, A., & Ibarra, C. (2021). Explaining the directionality of gang violence with court records. *Journal of Aggression, Conflict and Peace Research*, *13*(2/3), 83–95. https://doi.org/10.1108/JACPR-11-2020-0558

Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardeñes, J., Romance, M., Sendiña-Nadal, I., Wang, Z., & Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, *544*(1), 1–122. https://doi.org/10.1016/j.physrep.2014.07.001

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational and Mathematical Organization Theory*, *12*(1), 21–34. https://doi.org/10.1007/s10588-006-7084-x

Breuer, N., & Baradel, M. (2025). Investigating the dynamics of yakuza violence using multilevel network analysis. *Social Networks*, *82*, 14–26. https://doi.org/10.1016/j.socnet.2025.03.001.

Breuer, N., & Varese, F. (2023). The Structure of Trade-type and Governance-type Organized Crime Groups: A Network Study. *The British Journal of Criminology*, *63*(4), 867–888. https://doi.org/10.1093/bjc/azac065

Bright, D. A., Brewer, R., & Morselli, C. (2021). Using social network analysis to study crime: Navigating the challenges of criminal justice records. *Social Networks*, *66*, 50–64. https://doi.org/10.1016/j.socnet.2021.01.006

Bright, D. A., Greenhill, C., Britz, T., Ritter, A., & Morselli, C. (2017). Criminal network vulnerabilities and adaptations. *Global Crime*, *18*(4), 424–441. https://doi.org/10.1080/17440572.2017.1377614

Bright, D. A., Greenhill, C., & Levenkova, N. (2014). Dismantling Criminal Networks: Can Node Attributes Play a Role? In C. Morselli (Ed.), *Crime and Networks* (pp. 148–162). Routledge.

Bright, D. A., Greenhill, C., Ritter, A., & Morselli, C. (2015). Networks within networks: Using multiple link types to examine network structure and identify key actors in a drug trafficking operation. *Global Crime*, *16*(3), 219–237. https://doi.org/10.1080/17440572.2015.1039164

Bright, D. A., Koskinen, J., & Malm, A. (2019). Illicit Network Dynamics: The Formation and Evolution of a Drug Trafficking Network. *Journal of Quantitative Criminology*, *35*(2), 237–258. https://doi.org/10.1007/s10940-018-9379-8

Burt, R. S. (1995). *Structural Holes: The Social Structure of Competition*. Harvard University Press.

Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, *110*(2), 349–399. https://doi.org/10.1086/421787

Butts, C. T. (2008). Social network analysis: A methodological introduction. *Asian Journal Of Social Psychology*, *11*(1), 13–41. https://doi.org/10.1111/j.1467-839X.2007.00241.x

Calamunci, F., & Drago, F. (2020). The Economic Impact of Organized Crime Infiltration in the Legal Economy: Evidence from the Judicial Administration of Organized Crime Firms. *Italian Economic Journal*, *6*(2), 275–297. https://doi.org/10.1007/s40797-020-00128-x

Calamunci, F., Ferrante, L., & Scebba, R. (2022). Closed for mafia: Evidence from the removal of mafia firms on commercial property values. *Journal of Regional Science*, *62*(5), 1487–1511. https://doi.org/10.1111/jors.12612

Calderoni, F. (2012). The structure of drug trafficking mafias: The 'Ndrangheta and cocaine. *Crime, Law and Social Change*, *58*(3), 321–349. https://doi.org/10.1007/s10611-012-9387-9

Calderoni, F. (2014). Identifying Mafia Bosses from Meeting Attendance. In A. J. Masys (Ed.), *Networks and Network Analysis for Defence and Security* (pp. 27–48). Springer International Publishing. https://doi.org/10.1007/978-3-319-04147-6_2

Calderoni, F. (2016). Predicting Organized Crime Leaders. In G. Bichler & A. Malm (Eds.), *Disrupting Criminal Networks: Network Analysis in Crime Prevention* (pp. 89–110). Lynne Rienner Publishers. https://doi.org/10.1515/9781626372573

Calderoni, F., Brunetto, D., & Piccardi, C. (2017). Communities in criminal networks: A case study. *Social Networks*, *48*, 116–125. https://doi.org/10.1016/j.socnet.2016.08.003

Calderoni, F., & Superchi, E. (2019). The nature of organized crime leadership: Criminal leaders in meeting and wiretap networks. *Crime, Law and Social Change*, *72*(4), 419–444. https://doi.org/10.1007/s10611-019-09829-6

Campana, P. (2016). Explaining criminal networks: Strategies and potential pitfalls. *Methodological Innovations*, *9*, 1–10. https://doi.org/10.1177/2059799115622748

Campana, P., & Varese, F. (2013). Cooperation in criminal organizations: Kinship and violence as credible commitments. *Rationality and Society*, *25*(3), 263–289. https://doi.org/10.1177/1043463113481202

Campana, P., & Varese, F. (2022). Studying organized crime networks: Data sources, boundaries and the limits of structural measures. *Social Networks*, *69*, 149–159. https://doi.org/10.1016/j.socnet.2020.03.002

Caneppele, S., Calderoni, F., & Martocchia, S. (2009). Not only banks: Criminological models on the infiltration of public contracts by Italian organized crime. *Journal of Money Laundering Control*, *12*(2), 151–172. https://doi.org/10.1108/13685200910951910

Carley, K. M. (2015). Open-Source Exploitation for Understanding Covert Networks. In L. M. Gerdes (Ed.), *Illuminating Dark Networks* (1st ed., pp. 103–122). Cambridge University Press. https://doi.org/10.1017/CBO9781316212639.008

Catino, M. (2014). How Do Mafias Organize?: Conflict and Violence in Three Mafia Organizations. *European Journal of Sociology*, *55*(2), 177–220. https://doi.org/10.1017/S0003975614000095

Catino, M. (2019). *Mafia Organizations: The Visible Hand of Criminal Enterprise*. Cambridge University Press. https://doi.org/10.1017/9781108567183

Catino, M., Rocchi, S., & Vittucci Marzetti, G. (2022). The network of interfamily marriages in 'Ndrangheta. *Social Networks*, *68*, 318–329. https://doi.org/10.1016/j.socnet.2021.08.012

Cavallaro, L., Ficara, A., De Meo, P., Fiumara, G., Catanese, S., Bagdasar, O., Song, W., & Liotta, A. (2020). Disrupting resilient criminal networks through data analysis: The case of Sicilian Mafia. *PLOS ONE*, *15*(8), e0236476. https://doi.org/10.1371/journal.pone.0236476

Cellai, D., & Bianconi, G. (2016). Multiplex networks with heterogeneous activities of the nodes. *Physical Review E*, *93*(3), 032302. https://doi.org/10.1103/PhysRevE.93.032302

Chainey, S. P., & Alonso Berbotto, A. (2022). A structured methodical process for populating a crime script of organized crime activity using OSINT. *Trends in Organized Crime*, *25*(3), 272–300. https://doi.org/10.1007/s12117-021-09428-9

Chandra, S., Peters, S., & Zimmer, N. (2014). How Powdered Cocaine Flows Across the United States: Evidence From Open-Source Price Data. *Journal of Drug Issues*, *44*(4), 344–361. https://doi.org/10.1177/0022042614522621

Checchi, V. V., & Polo, M. (2020). Blowing in the Wind: The Infiltration of Sicilian Mafia in the Wind Power Business. *Italian Economic Journal*, *6*(2), 325–353. https://doi.org/10.1007/s40797-020-00126-z

Cheloukhine, S. (2008). The roots of Russian organized crime: From old-fashioned professionals to the organized criminal groups of today. *Crime, Law and Social Change*, *50*(4–5), 353–374. https://doi.org/10.1007/s10611-008-9117-5

Chen, Y., Paul, G., Cohen, R., Havlin, S., Borgatti, S. P., Liljeros, F., & Eugene Stanley, H. (2007). Percolation theory and fragmentation measures in social networks. *Physica A: Statistical Mechanics and Its Applications*, *378*(1), 11–19. https://doi.org/10.1016/j.physa.2006.11.074

Chu, Y. K. (2000). *The Triads as Business*. Taylor & Francis Group.

Coutinho, J. A., Diviák, T., Bright, D., & Koskinen, J. (2020). Multilevel determinants of collaboration between organised criminal groups. *Social Networks*, *63*, 56–69. https://doi.org/10.1016/j.socnet.2020.04.002

Cowley, A., Ryan, C. M., & Dunn, E. C. (2015). The Law, the Mafia, and the Production of Sovereignties in the Kyrgyz Penal System. *Ab Imperio*, *2015*(2), 183–208. https://doi.org/10.1353/imp.2015.0048

Davis, G. F. (1996). The Significance of Board Interlocks for Corporate Governance. *Corporate Governance: An International Review*, *4*(3), 154–159. https://doi.org/10.1111/j.1467-8683.1996.tb00144.x

Davis, G. F., Yoo, M., & Baker, W. E. (2003). The Small World of the American Corporate Elite, 1982-2001. *Strategic Organization*, *1*(3), 301–326. https://doi.org/10.1177/14761270030013002

De Andrade, R. L., Rêgo, L. C., Coelho Da Silva, T. L., De Macêdo, J. A. F., & Silva, W. C. P. (2021). Energy disruptive centrality with an application to criminal network. *Communications in Nonlinear Science and Numerical Simulation*, *99*, 105834. https://doi.org/10.1016/j.cnsns.2021.105834

De La Mora Tostado, S., Hernández-Vargas, E. A., & Núñez-López, M. (2024). Modeling human trafficking and the limits of dismantling strategies. *Social Network Analysis and Mining*, *14*(1), 84. https://doi.org/10.1007/s13278-024-01208-x

Decker, S. H., & Curry, G. D. (2002). Gangs, gang homicides, and gang loyalty: Organized crimes or disorganized criminals. *Journal of Criminal Justice*, *30*(4), 343–352. https://doi.org/10.1016/S0047-2352(02)00134-4

Decker, S. H., Katz, C. M., & Webb, V. J. (2008). Understanding the Black Box of Gang Organization: Implications for Involvement in Violent Crime, Drug Sales, and Violent Victimization. *Crime & Delinquency*, *54*(1), 153–172. https://doi.org/10.1177/0011128706296664

Dekker, D., Krackhardt, D., & Snijders, T. A. B. (2007). Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions. *Psychometrika*, *72*(4), 563–581. https://doi.org/doi:10.1007/s11336-007-9016-1

DellaPosta, D. (2017). Network closure and integration in the mid-20th century American mafia. *Social Networks*, *51*, 148–157. https://doi.org/10.1016/j.socnet.2016.11.005

DellaPosta, D. (2023). Bridging the parochial divide: Outsider brokerage in mafia families. *Social Science Research*, *114*, 102913. https://doi.org/10.1016/j.ssresearch.2023.102913

Descormiers, K., & Morselli, C. (2011). Alliances, Conflicts, and Contradictions in Montreal's Street Gang Landscape. *International Criminal Justice Review*, *21*(3), 297–314. https://doi.org/10.1177/1057567711418501

DIA. (2020). *Relazione del Ministro dell'Interno al Parlamento. Attività Svolta e Risultati Conseguiti (Luglio-Dicembre)*. Direzione Investigativa Antimafia. https://direzioneinvestigativaantimafia.interno.gov.it/semestrali/sem/2020/2sem2020.pdf

Diviák, T. (2023). Structural resilience and recovery of a criminal network after disruption: A simulation study. *Journal of Experimental Criminology*. https://doi.org/10.1007/s11292-023-09563-z

Diviák, T., Dijkstra, J. K., & Snijders, T. A. B. (2019). Structure, multiplexity, and centrality in a corruption network: The Czech Rath affair. *Trends in Organized Crime*, *22*(3), 274–297. https://doi.org/10.1007/s12117-018-9334-y

Diviák, T., Dijkstra, J. K., van der Wijk, F., Oosting, I., & Wolters, G. (2021). Women trafficking networks: Structure and stages of women trafficking in five Dutch small-scale networks. *European Journal of Criminology*, *20*(4), 1506–1528. https://doi.org/10.1177/14773708211053135

Diviák, T., & Lord, N. (2024). Connecting the Corrupt: Data Sources to Study Networks of Serious Financial Crime in the United Kingdom. In L. Huey & D. Buil-Gil (Eds.), *The Crime Data Handbook* (1st ed., pp. 152–165). Bristol University Press. https://www.jstor.org/stable/jj.9692690.17

Diviák, T., Van Nassau, C. S., Dijkstra, J. K., & Snijders, T. A. B. (2022). Dynamics and disruption: Structural and individual changes in two Dutch Jihadi networks after police interventions. *Social Networks*, *70*, 364–374. https://doi.org/10.1016/j.socnet.2022.04.001

Dolliver, D. S., Ericson, S. P., & Love, K. L. (2018). A Geographic Analysis of Drug Trafficking Patterns on the TOR Network. *Geographical Review*, *108*(1), 45–68. https://doi.org/10.1111/gere.12241

Duijn, P. A. C., Kashirin, V., & Sloot, P. M. A. (2014). The Relative Ineffectiveness of Criminal Network Disruption. *Scientific Reports*, *4*(1), 4238. https://doi.org/10.1038/srep04238

Europol. (2021). *European Union Serious and Organised Crime Threat Assessment (SOCTA) 2021*. https://www.europol.europa.eu/publication-events/main-reports/european-union-serious-and-organised-crime-threat-assessment-socta-2021

Falzon, L., Quintane, E., Dunn, J., & Robins, G. (2018). Embedding time in positions: Temporal measures of centrality for social network analysis. *Social Networks*, *54*, 168–178. https://doi.org/10.1016/j.socnet.2018.02.002

Ficara, A., Curreri, F., Fiumara, G., & De Meo, P. (2022). Human and social capital strategies for Mafia network disruption. *CoRR*, *abs/2209.02012*. https://doi.org/10.48550/arXiv.2209.02012

Ficara, A., Curreri, F., Fiumara, G., De Meo, P., & Liotta, A. (2022). Covert Network Construction, Disruption, and Resilience: A Survey. *Mathematics*, *10*(16), 2929. https://doi.org/10.3390/math10162929

Ficara, A., Fiumara, G., De Meo, P., & Catanese, S. (2021). Multilayer Network Analysis: The Identification of Key Actors in a Sicilian Mafia Operation. In D. Perakovic & L. Knapcikova (Eds.), *Future Access Enablers for Ubiquitous and Intelligent Infrastructures* (Vol. 382, pp. 120–134). Springer International Publishing. https://doi.org/10.1007/978-3-030-78459-1_9

Friedkin, N. E. (1983). Horizons of Observability and Limits of Informal Control in Organizations. *Social Forces*, *62*(1), 54–77.

Frisby, T. (1998). The rise of organised crime in Russia: Its roots and social significance. *Europe-Asia Studies*, *50*(1), 27–49. https://doi.org/10.1080/09668139808412522

Galeotti, M. (2004). The Russian "Mafiya": Consolidation and Globalisation. *Global Crime*, *6*(1), 54–69. https://doi.org/10.1080/1744057042000297972

Galeotti, M. (2018). *The Vory: Russia's Super Mafia*. Yale University Press.

Gambetta, D. (1993). *The Sicilian Mafia: The business of private protection*. Harvard University Press.

Garcia-Bernardo, J., Fichtner, J., Takes, F. W., & Heemskerk, E. M. (2017). Uncovering Offshore Financial Centers: Conduits and Sinks in the Global Corporate Ownership Network. *Scientific Reports*, *7*(1), 6246. https://doi.org/10.1038/s41598-017-06322-9

Gennaioli, C., & Tavoni, M. (2016). Clean or dirty energy: Evidence of corruption in the renewable energy sector. *Public Choice*, *166*(3–4), 261–290. https://doi.org/10.1007/s11127-016-0322-y

Gilinskiy, Y., & Kostjukovsky, Y. (2004). From Thievish Artel to Criminal Corporation: The History of Organised Crime in Russia. In C. Fijnaut & L. Paoli (Eds.), *Organised Crime in Europe* (pp. 181–202). Springer International Publishing. https://doi.org/10.1007/978-1-4020-2765-9_6

Glattfelder, J. B., & Battiston, S. (2009). Backbone of complex networks of corporations: The flow of control. *Physical Review E*, *80*(3), 036104. https://doi.org/10.1103/PhysRevE.80.036104

Gragert, L. B. A. (1997). Yakuza: The Warlords of Japanese Organized Crime. *Annual Survey of International & Comparative Law*, *4*(1), 147–204.

Gravel, J., Allison, B., West-Fagan, J., McBride, M., & Tita, G. E. (2018). Birds of a Feather Fight Together: Status-Enhancing Violence, Social Distance and the Emergence of Homogenous Gangs. *Journal of Quantitative Criminology*, *34*(1), 189–219. https://doi.org/10.1007/s10940-016-9331-8

Gurciullo, S. (2014). *Organised crime's infiltration in the legitimate private economy: An empirical network analysis approach*. arXiv. https://doi.org/10.48550/arXiv.1403.5071

Haller, M. H. (1990). Illegal Enterprise: A Theoretical and Historical Interpretation. *Criminology*, *28*(2), 207–236. https://doi.org/10.1111/j.1745-9125.1990.tb01324.x

Handelman, S. (1995). *Comrade Criminal: Russia's New Mafiya*. Yale University Press.

Hill, P. B. E. (2003). *The Japanese Mafia: Yakuza, Law, and the State*. Oxford University Press. https://doi.org/10.1093/0199257523.001.0001

Hollway, J., & Koskinen, J. (2016). Multilevel embeddedness: The case of the global fisheries governance complex. *Social Networks*, *44*, 281–294. https://doi.org/10.1016/j.socnet.2015.03.001

Hughes, L. A., & Short, J. F. (2005). Disputes Involving Youth Street Gang Members: Micro-Social Contexts. *Criminology*, *43*(1), 43–76. https://doi.org/10.1111/j.0011-1348.2005.00002.x

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008). Goodness of Fit of Social Network Models. *Journal of the American Statistical Association*, *103*(481), 248–258. https://doi.org/10.1198/016214507000000446

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *Journal of Statistical Software*, *24*(3), 1–29. https://doi.org/10.18637/jss.v024.i03

InfoCamere. (n.d.). *Visura Camerale Ordinaria e Storica e Certificato*. Il Registro Imprese. https://www.registroimprese.it/visura-camerale-ordinaria-storica-e-certificato

Jacobs, J. B. (2020). The Rise and Fall of Organized Crime in the United States. *Crime and Justice*, *49*(1), 17–67. https://doi.org/10.1086/706895

Jahanpour, E., & Chen, X. (2013). Analysis of complex network performance and heuristic node removal strategies. *Communications in Nonlinear Science and Numerical Simulation*, *18*(12), 3458–3468. https://doi.org/10.1016/j.cnsns.2013.04.030

Jia, M., De Meo, P., Gabrys, B., & Musial, K. (2024). Network disruption via continuous batch removal: The case of Sicilian Mafia. *PLOS ONE*, *19*(8), e0308722. https://doi.org/10.1371/journal.pone.0308722

Jitsuwa Jidai Henshūbuhen. (2002). *Yamaguchi-gumi Daikaibō: Zenkoku Shinkō to Sanman'nin Taisei no Ishizue [The Anatomy of the Yamaguchi-gumi: Nationwide invasion and the foundation of the 30,000 man system]*. Sanwa Shuppan.

Jitsuwa Jidai Henshūbuhen. (2007a). *Gendai Yakuza Daijiten [Modern Yakuza Encyclopedia]*. Yosensha.

Jitsuwa Jidai Henshūbuhen. (2007b). *Yamaguchi-gumi Sandaime: Seiha no Yabō [Third-generation Yamaguchi-gumi: Ambition to Conquer]*. Sanwa Shuppan.

Jofre, M. (2022). Network analysis for financial crime risk assessment: The case study of the gambling division in Malta. *Global Crime*, *23*(2), 148–170. https://doi.org/10.1080/17440572.2022.2077330

Jofre, M., Bosisio, A., & Riccardi, M. (2024). Financial crime risk assessment: Machine learning insights into ownership structures in secrecy firms. *Global Crime*, 1–26. https://doi.org/10.1080/17440572.2024.2402848

Kaplan, D. E., & Dubro, A. (2012). *Yakuza: Japan's criminal underworld* (25th anniversary ed.). University of California Press.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203–271. https://doi.org/10.1093/comnet/cnu016

Kogut, B., & Walker, G. (2001). The Small World of Germany and the Durability of National Networks. *American Sociological Review*, *66*(3), 317–335. https://doi.org/10.1177/000312240106600301

Kojaku, S., & Masuda, N. (2017). Finding multiple core-periphery pairs in networks. *Physical Review E*, *96*(5), 052313. https://doi.org/10.1103/PhysRevE.96.052313

Kojaku, S., & Masuda, N. (2018). Core-periphery structure requires something else in the network. *New Journal of Physics*, *20*(4), 043012. https://doi.org/10.1088/1367-2630/aab547

Krackhardt, D. (1988). Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data. *Social Networks*, *10*(4), 359–381. https://doi.org/10.1016/0378-8733(88)90004-4

Krajewski, A. T., DellaPosta, D., & Felmlee, D. (2022). Vertical organizations, flat networks: Centrality and criminal collaboration in the Italian-American Mafia. *Social Networks*, *68*, 127–138. https://doi.org/10.1016/j.socnet.2021.06.001

Krebs, V. E. (2002). Mapping Networks of Terrorist Cells. *Connections*, *24*(3), 43–52.

Kruisbergen, E. W., Kleemans, E. R., & Kouwenberg, R. F. (2015). Profitability, Power, or Proximity? Organized Crime Offenders Investing Their Money in Legal Economy. *European Journal on Criminal Policy and Research*, *21*(2), 237–256. https://doi.org/10.1007/s10610-014-9263-5

Kupatadze, A. (2014). Prisons, politics and organized crime: The case of Kyrgyzstan. *Trends in Organized Crime*, *17*(3), 141–160. https://doi.org/10.1007/s12117-013-9208-2

Lazega, E., Jourda, M.-T., Mounier, L., & Stofer, R. (2008). Catching up with big fish in the big pond? Multi-level network analysis through linked design. *Social Networks*, *30*(2), 159–176. https://doi.org/10.1016/j.socnet.2008.02.001

Levy, M. (2016a). gwdegree: Improving interpretation of geometrically-weighted degree estimates in exponential random graph models. *The Journal of Open Source Software*, *1*(3), 36. https://doi.org/10.21105/joss.00036

Levy, M. (2016b, June). *Interpretation of GW-Degree Estimates in ERGMs* [Poster]. Political Networks Conference. https://doi.org/10.6084/m9.figshare.3465020.v1

Lewis, K., & Papachristos, A. V. (2020). Rules of the Game: Exponential Random Graph Models of a Gang Homicide Network. *Social Forces*, *98*(4), 1829–1858. https://doi.org/10.1093/sf/soz106

Lomi, A., Robins, G., & Tranmer, M. (2016). Introduction to multilevel social networks. *Social Networks*, *44*, 266–268. https://doi.org/10.1016/j.socnet.2015.10.006

Lusher, D., Koskinen, J., & Robins, G. (Eds.). (2012). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511894701

Malm, A., Bichler, G., & Van De Walle, S. (2010). Comparing the ties that bind criminal networks: Is blood thicker than water? *Security Journal*, *23*(1), 52–74. https://doi.org/10.1057/sj.2009.18

Manzi, D., & Calderoni, F. (2024a). An Agent-Based Model for Assessing the Resilience of Drug Trafficking Organizations to Law Enforcement Interventions. *Journal of Artificial Societies and Social Simulation*, *27*(3), 3. https://doi.org/10.18564/jasss.5430

Manzi, D., & Calderoni, F. (2024b). The resilience of drug trafficking organizations: Simulating the impact of police arresting key roles. *Journal of Criminal Justice*, *91*, 102165. https://doi.org/10.1016/j.jcrimjus.2024.102165

Massari, M., & Martone, V. (2018). Doing Research on Mafia Violence. In M. Massari & V. Martone (Eds.), *Mafia Violence: Political, Symbolic, and Economic Forms of Violence in Camorra Clans* (pp. 1–16). Routledge. https://doi.org/10.4324/9780429467554

Mastrobuoni, G. (2015). The Value of Connections: Evidence from the Italian-American Mafia. *The Economic Journal*, *125*(586), F256–F288. https://doi.org/10.1111/ecoj.12234

Mastrobuoni, G., & Patacchini, E. (2012). Organized Crime Networks: An Application of Network Analysis Techniques to the American Mafia. *Review of Network Economics*, *11*(3). https://doi.org/10.1515/1446-9022.1324

Mirenda, L., Mocetti, S., & Rizzica, L. (2019). *The real effect of ’ndrangheta: Firm-level evidence* (Working Paper 1235). Bank of Italy.

Mironova, V. (2023). *Criminals, Nazis, and Islamists: Competition for Power in Former Soviet Union Prisons* (1st ed.). Oxford University PressNew York. https://doi.org/10.1093/oso/9780197645659.003.0011

Mizoguchi, A., & Suzuki, T. (2021). *Shokugyō to shite no yakuza [Yakuza as a profession]*. Kogakukan Shinsho.

Moro, F. N., & Sberna, S. (2018). Transferring Violence? Mafia Killings in Nontraditional Areas: Evidence from Italy. *Journal of Conflict Resolution*, *62*(7), 1579–1601. https://doi.org/10.1177/0022002717693049

Morselli, C. (2001). Structuring Mr. Nice: Entrepreneurial opportunities and brokerage positioning in the cannabis trade. *Crime, Law and Social Change*, *35*, 203–244. https://doi.org/10.1023/A:1011272411727

Morselli, C. (2003). Career opportunities and network-based privileges in the Cosa Nostra. *Crime, Law and Social Change*, *39*, 383–418. https://doi.org/10.1023/A:1024020609694

Morselli, C. (2010). Assessing Vulnerable and Strategic Positions in a Criminal Network. *Journal of Contemporary Criminal Justice*, *26*(4), 382–392. https://doi.org/10.1177/1043986210377105

Morselli, C., & Petit, K. (2007). Law-Enforcement Disruption of a Drug Importation Network. *Global Crime*, *8*(2), 109–130. https://doi.org/10.1080/17440570701362208

Morselli, C., & Roy, J. (2008). Brokerage Qualifications in Ringing Operations. *Criminology*, *46*(1), 71–98. https://doi.org/10.1111/j.1745-9125.2008.00103.x

Musciotto, F., & Miccichè, S. (2022). Effective strategies for targeted attacks to the network of Cosa Nostra affiliates. *EPJ Data Science*, *11*(1), 1–14. https://doi.org/10.1140/epjds/s13688-022-00323-z

Musotto, R. (2022). *Understanding Mafia Networks* (1st ed.). Routledge. https://doi.org/10.4324/9780429203206

Nakamura, K., Tita, G. E., & Krackhardt, D. (2020). Violence in the “balance”: A structural analysis of how rivals, allies, and third-parties shape inter-gang violence. *Global Crime*, *21*(1), 3–27. https://doi.org/10.1080/17440572.2019.1627879

Natarajan, M. (2000). Understanding the Structure of a Drug Trafficking Organization: A Conversational Analysis. *Crime Prevention Studies*, *11*, 273–298.

Natarajan, M. (2006). Understanding the Structure of a Large Heroin Distribution Network: A Quantitative Analysis of Qualitative Data. *Journal of Quantitative Criminology*, *22*(2), 171–192. https://doi.org/10.1007/s10940-006-9007-x

Neumann, M., Lotzmann, U., & Troitzsch, K. G. (2017). Mafia war: Simulating conflict escalation in criminal organizations. *Trends in Organized Crime*, *20*(1–2), 139–178. https://doi.org/10.1007/s12117-016-9292-1

Newman, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, *64*(1), 016132. https://doi.org/10.1103/PhysRevE.64.016132

NPA. (2007). *Tokushū: Soshiki Hanzai wo Yurusanai Shakai wo Mezashite [Special Report: Aiming for a Society that does not Tolerate Organised Crime]*. National Police Agency. https://www.npa.go.jp/hakusyo/h19/honbun/html/jca00000.html

NPA. (2012). *Keisatsu Hakusho [White Paper of the Police]*. National Police Agency. https://www.npa.go.jp/hakusyo/h24/index.html

NPA. (2022). *Keisatsu Hakusho [White Paper of the Police]*. National Police Agency. https://www.npa.go.jp/hakusyo/r04/index.html

Oliver, K., Crossley, N., Edwards, G., Koskinen, J., & Everett, M. (2014). *Covert networks: Structures, processes and types* (pp. 1–24) [Working paper]. The Mitchell Centre for Social Network Analysis, University of Manchester. https://www.socialsciences.manchester.ac.uk/mitchell-centre/research/covert-networks/outputs/

Open Knowledge Foundation. (n.d.). *What is Open Data?* Open Data Handbook. https://opendatahandbook.org/guide/en/what-is-open-data/

Opsahl, T. (2013). Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks*, *35*(2), 159–167. https://doi.org/10.1016/j.socnet.2011.07.001

Opsahl, T., Colizza, V., Panzarasa, P., & Ramasco, J. J. (2008). Prominence and Control: The Weighted Rich-Club Effect. *Physical Review Letters*, *101*(16), 168702. https://doi.org/10.1103/PhysRevLett.101.168702

P. B., D., Johnson, T. P., & Balakrishnan, K. (2023). A study of the effectiveness of the Profile Closeness Attack on the Sicilian Mafia Network. *2023 4th International Conference on Innovative Trends in Information Technology (ICITIIT)*, 1–6. https://doi.org/10.1109/ICITIIT57246.2023.10068719

Paoli, L. (2004). Italian Organised Crime: Mafia Associations and Criminal Enterprises. *Global Crime*, *6*(1), 19–31. https://doi.org/10.1080/1744057042000297954

Paoli, L. (2020). What Makes Mafias Different? *Crime and Justice*, *49*, 141–222. https://doi.org/10.1086/708826

Papachristos, A. V. (2009). Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. *American Journal of Sociology*, *115*(1), 74–128. https://doi.org/10.1086/597791

Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence. *American Sociological Review*, *78*(3), 417–447. https://doi.org/10.1177/0003122413486800

Pinto Hidalgo, J. J., & Silva Centeno, J. A. (2022). Geospatial Intelligence and Artificial Intelligence for Detecting Potential Coca Paste Production Infrastructure in the Border Region of Venezuela and Colombia. *Journal of Applied Security Research*, *18*(4), 1000–1050. https://doi.org/10.1080/19361610.2022.2111184

PrimeCrime.ru. (n.d.). *Lyudi*. PrimeCrime.Ru. https://www.primecrime.ru/characters/

PrimeCrime.ru. (2013, April 25). Nika za glavnuyu rol': V Moskve zaderzhan «vor v zakone» Nika Gagrinskii. *PrimeCrime.ru*. https://www.primecrime.ru/news/2013-04-25_5440

PrimeCrime.ru. (2015, January 6). "Zhizn" vor.am!: V Armenii prokhodit istoricheskaya vorovskaya skhodka'. *PrimeCrime.Ru*. http://www.primecrime.ru/news/2015-01-06_5894/

PrimeCrime.ru. (2020, March 9). "'Dlya menya oni – ne vory!": Na skhodke, organizovannoi Guli, Vazha Biganishvili kosnulsya chesti desyati vorov'. *PrimeCrime.Ru*. https://www.primecrime.ru/news/2020-03-09_7425/

Randle, J., & Bichler, G. (2017). Uncovering the Social Pecking Order in Gang Violence. In B. LeClerc & E. U. Savona (Eds.), *Crime Prevention in the 21st Century* (pp. 165–186). Springer International Publishing. https://doi.org/10.1007/978-3-319-27793-6_12

Reuter, P. (1983). *Disorganized Crime: The Economics of the Visible Hand*. MIT Press.

Reuter, P., & Paoli, L. (2020). How Similar Are Modern Criminal Syndicates to Traditional Mafias? *Crime and Justice*, *49*, 223–287. https://doi.org/10.1086/708869

Riccardi, M. (2014). When Criminals Invest in Businesses: Are We Looking in the Right Direction? An Exploratory Analysis of Companies Controlled by Mafias. In S. Caneppele & F. Calderoni (Eds.), *Organized Crime, Corruption and Crime Prevention: Essays in Honor of Ernesto U. Savona* (pp. 197–206). Springer International Publishing. https://doi.org/10.1007/978-3-319-01839-3

Riccardi, M., & Berlusconi, G. (2016). Measuring organised crime infiltration in legal businesses. In E. U. Savona, M. Riccardi, & G. Berlusconi (Eds.), *Organised Crime in European Businesses* (pp. 16–32). Routledge. https://doi-org.ezproxy-prd.bodleian.ox.ac.uk/10.4324/9781315640617

Riccardi, M., & Maggioni, M. A. (2025). Bricks or cooks? Geographical and social determinants of the investment choices of mafia-type organized crime. *Criminology & Criminal Justice*, *25*(1), 171–199. https://doi.org/10.1177/17488958241293927

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 173–191. https://doi.org/10.1016/j.socnet.2006.08.002

Savona, E. U. (2015). The Businesses of Italian Mafias. *European Journal on Criminal Policy and Research*, *21*(2), 217–236. https://doi.org/10.1007/s10610-014-9261-7

Schelling, T. C. (1980). *The Strategy of Conflict*. Harvard University Press.

Schreck, C. (2009, August 5). The Who's Who of Moscow mobsters. *The National*. https://www.thenationalnews.com/world/europe/the-who-s-who-of-moscow-mobsters-1.493579

Sciarrone, R., & Storti, L. (2014). The territorial expansion of mafia-type organized crime. The case of the Italian mafia in Germany. *Crime, Law and Social Change*, *61*(1), 37–60. https://doi.org/10.1007/s10611-013-9473-7

Serio, J. D., & Razinkin, V. (1995). Thieves Professing the Code: The Traditional Role of the Vory v Zakone in Russia's Criminal Underworld. *Low Intensity Conflict & Law Enforcement*, *4*(1), 72–88.

Shavitt, Y., & Singer, Y. (2007). Beyond centrality—Classifying topological significance using backup efficiency and alternative paths. *New Journal of Physics*, *9*(8), 266–266. https://doi.org/10.1088/1367-2630/9/8/266

Shortland, A., & Varese, F. (2016). State-Building, Informal Governance and Organised Crime: The Case of Somali Piracy. *Political Studies*, *64*(4), 811–831. https://doi.org/10.1111/1467-9248.12227

Slade, G. (2014). *Reorganizing Crime: Mafia and Anti-Mafia in Post-Soviet Georgia*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199674640.001.0001

Slade, G. (2016). Violence as Information During Prison Reform: Evidence From the Post-Soviet Region. *British Journal of Criminology*, *56*(5), 937–955. https://doi.org/10.1093/bjc/azv085

Slade, G., & Azbel, L. (2020). Managing drugs in the prisoner society: Heroin and social order in Kyrgyzstan's prisons. *Punishment & Society*, *24*(1), 26–45. https://doi.org/10.1177/1462474520956280

Smith, C. M., & Papachristos, A. V. (2016). Trust Thy Crooked Neighbor: Multiplexity in Chicago Organized Crime Networks. *American Sociological Review*, *81*(4), 644–667. https://doi.org/10.1177/0003122416650149

Smith, D. C. (1971). Some Things that May be More Important to Understand About Organized Crime than Costa Nostra. *University of Florida Law Review*, *24*(1), 1–31.

Snijders, T. A. B., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New Specifications for Exponential Random Graph Models. *Sociological Methodology*, *36*(99–154).

Stadtfeld, C., & Amati, V. (2021). Network mechanisms and network models. In G. Manzo (Ed.), *Research Handbook on Analytical Sociology* (pp. 432–452). Edward Elgar Publishing. https://doi.org/10.4337/9781789906851.00032

Stys, P., Verweijen, J., Muzuri, P., Muhindo, S., Vogel, C., & Koskinen, J. H. (2020). Brokering between (not so) overt and (not so) covert networks in conflict zones. *Global Crime*, *21*(1), 74–110. https://doi.org/10.1080/17440572.2019.1596806

Thompson, H. A. (2024). The Industrial Organization of the Mafia. *Journal of Law and Economics*, 555–597.

Tita, G. E., & Radil, S. M. (2011). Spatializing the Social Networks of Gangs to Explore Patterns of Violence. *Journal of Quantitative Criminology*, *27*(4), 521–545. https://doi.org/10.1007/s10940-011-9136-8

Toledo, A. S. O., Carpi, L. C., & Atman, A. P. F. (2020). Diversity Analysis Exposes Unexpected Key Roles in Multiplex Crime Networks. In H. Barbosa, J. Gomez-Gardenes, B. Gonçalves, G. Mangioni, R. Menezes, & M. Oliveira (Eds.), *Complex Networks XI: Proceedings of the 11th Conference on Complex Networks CompleNet 2020* (pp. 371–382). Springer International Publishing. https://doi.org/10.1007/978-3-030-40943-2

Toledo, A. S. O., Carpi, L. C., Atman, A. P. F., & Scarpelli, A. P. B. (2023). Multiplex key roles to disrupt criminal networks. *Social Network Analysis and Mining*, *13*(1), 98. https://doi.org/10.1007/s13278-023-01101-z

Tsereteli, I. (2022). Some Historical Aspects of the Establishment and Development of the "Vor V Zakone" E Bis Institute. *Journal of Legal Studies*, *29*(43), 42–53. https://doi.org/10.2478/jles-2022-0003

Tumminello, M., Petruzzella, F., Ferrara, C., & Miccichè, S. (2021). Anagraphical relationships and crime specialization within Cosa Nostra. *Social Networks*, *64*, 29–41. https://doi.org/10.1016/j.socnet.2020.07.011

Varese, F. (2001). *The Russian Mafia: Private Protection in a New Market Economy*. Oxford University Press. https://hdl.handle.net/2027/heb32498.0001.001

Varese, F. (2013a). Protection and Extortion. In L. Paoli (Ed.), *The Oxford Handbook of Organized Crime*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199730445.013.020

Varese, F. (2013b). The Structure and the Content of Criminal Connections: The Russian Mafia in Italy. *European Sociological Review*, *29*(5), 899–909. https://doi.org/10.1093/esr/jcs067

Varese, F. (2017). *Mafia Life: Love, Death, and Money at the Heart of Organized Crime*. Profile Books.

Varese, F. (2024). *Russia in Four Criminals*. Polity.

Varese, F., Lonsky, J., & Podvysotskiy, Y. (2021). The Resilience of the Russian mafia: An Empirical Study. *The British Journal of Criminology*, *61*(1), 143–166. https://doi.org/10.1093/bjc/azaa053

Vershov, Y. (2018, December 2). «Vor v zakone» Kostyl' poluchil svobodu i problemy. *Rosbalt News Agency*. https://www.rosbalt.ru/moscow/2018/12/02/1750183.html

Villani, S., Mosca, M., & Castiello, M. (2019). A virtuous combination of structural and skill analysis to defeat organized crime. *Socio-Economic Planning Sciences*, *65*, 51–65. https://doi.org/10.1016/j.seps.2018.01.002

Wang, P., Robins, G., Pattison, P., & Lazega, E. (2013). Exponential random graph models for multilevel networks. *Social Networks*, *35*(1), 96–115. https://doi.org/10.1016/j.socnet.2013.01.004

Wang, P., Sharpe, K., Robins, G., & Pattison, P. E. (2009). Exponential random graph (p*) models for affiliation networks. *Social Networks*, *31*(1), 12–25. https://doi.org/10.1016/j.socnet.2008.08.002

Wood, G. (2017). The structure and vulnerability of a drug trafficking collaboration network. *Social Networks*, *48*, 1–9. https://doi.org/10.1016/j.socnet.2016.07.001

Yeung, W. N., Clemente, R. D., & Lambiotte, R. (2025). *Garbage in Garbage out: Impacts of data quality on criminal network intervention* (arXiv:2501.01508). arXiv. https://doi.org/10.48550/arXiv.2501.01508

Zhang, S., & Chin, K. (2003). The Declining Significance of Triad Societies in Transnational Illegal Activities: A Structural Deficiency Perspective. *The British Journal of Criminology*, *43*(3), 469–488. https://doi.org/10.1093/bjc/43.3.469