Uneven Development and Insurgency in Turkey:
A Computational Approach

Ollie Ballinger

St. Anne's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2022

to Mom, Dad, and Sean.

Acknowlegements

I owe a debt of gratitude to a number of people without whom this work would not have been possible, and to whom I'd like to extend my deepest thanks. First and foremost, my supervisor Adeel Malik for being there every step of the way and providing me with invaluable guidance on everything from historical context to standard errors; Doug Gollin, whose mentorship, advice, and tireless support shaped my thinking, honed my skills, and deepened my interest in understanding the human world; Bernie Hogan, for helping me navigate an unfamiliar field and prompting me to question the nature of my data and the inferences that could be drawn from it; Günes Murat Tezcür, whose work provided a starting point for my own and whose early advice shaped the trajectory of my research; Arun Frey, Caliean Osborne, Felix Soliman, and Neave O'Clery for reading my drafts and providing me with thorough and thoughtful feedback; Sean, for getting me from hello world to neural networks; Mom, Dad, Dede, and AnneAnne for the love, support, and patliçan kizartmasi; Melissa and Nick, who became family when the world was turned upside down; and Emma, without whom this journey would never have even started.

Abstract

This thesis develops computational techniques to gather, process, and analyze fine-grained data on the war between the Turkish State and the *Partiya Karkerên Kurdistan* (PKK), a Kurdish insurgent group. In three chapters, I seek to better understand the reasons that lead people join the PKK, assess the impact of development policies aimed at dissuading them from doing so, and explain the group's structural resilience to military force. The first chapter explores the recruitment of young urban recruits using web scraping, fuzzy matching, and other computational techniques. Leveraging an unprecedentedly detailed research design, militants are compared to random samples of Turkish citizens. I find evidence linking insurgent recruitment and a range factors including birth order and family size, peer effects, and conflict-induced migration. The second chapter explores economic motivations in Turkey's agrarian Southeast using remote sensing and spatial econometrics. I find that clashes are more frequent following poor harvests, irrigation decouples agricultural income from rainfall, and that conflict was reduced in areas benefiting from a state-sponsored agricultural development program. The third chapter examines the PKK as a whole, focusing on its structural characteristics. I develop a new methodology that leverages deep learning to create a social network graph based on co-appearance in photographs which retains many of the broad structural features of the PKK. Analytical results indicate that the densely interconnected nature of the PKK makes it highly robust to a range of counterinsurgency strategies. Together, substantive findings suggests that development policy is a far more promising avenue for the resolution of the conflict than military policy, while the methodological contributions include the development of forward looking analytical techniques and open source software that enable highly detailed quantitative analysis of civil conflict.

Contents

List of Figures

Introduction

The woman above is Gülnaz Ekinci. She was born in 1984, in the city of Konya, Turkey. Her parents moved there from a farming village in the country's arid Southeastern region, which receives as much rainfall as parts of the Sahara Desert. Like many others, they left their family and friends behind to pursue new opportunities in Turkey's rapidly growing cities. The rest of her family– members of the Kurdish ethnic minority– defied the Turkish state's attempts at assimilation. Six months after Gülnaz was born, Kurdish resistance to these measures coalesced into a more organized form: the foundation of the Kurdistan Workers' Party (Paritya Karkeren Kürdistane, henceforth PKK), an insurgent group that has been engaged in armed struggle against the Turkish state ever since. Members of Gülnaz's family, still in the mostly Kurdish Southeast, supported the rebels. During visits to her family's village throughout her childhood, she met several PKK Guerillas who left a lasting impression. Despite her Kurdish heritage, having grown up in a largely Turkish city, she never learned her parents' language. In a first act of resistance, she began learning Kurdish. Upon her return to Konya, she took a further step: along with friends and family, she began volunteering for the political wing of the PKK. Finally, at the age of 15, Gülnaz Ekinci traveled more than 1,400 kilometers east to the mountains of Iran where she began training as a guerilla fighter with an all-female unit of the PKK. Over the course of more than a decade, she participated in numerous operations against Turkish forces, was captured and escaped, and eventually rose to the rank of Commander. On February 13th 2012– the day after her birthday– she was killed in the village of Besta near the Turkish border with Iraq.

This profile of Gülnaz Ekinci was not put together through interviews, field research, or ethnography. It was constructed entirely from open source, public domain material that anyone with an internet connection can access; some of it was published online by the PKK itself, some of it was buried in leaked government data, some of it was drawn from earth-observing satellites. Gülnaz is not unique in this regard: similarly detailed profiles can be constructed for thousands of deceased PKK militants. By using computational research methods to triangulate open source data, we can begin to uncover the life stories of people like Gulnaz Ekinci and in doing so, better understand her journey to insurgency and the junctures that may have averted it: Would her parents have moved to Konya had opportunities been more plentiful on their farm? What role did her social network in Konya play in her decision to join the PKK? And how, despite the killing of thousands of people like Gülnaz Ekinci, has the PKK managed to survive? Though the details of Gülnaz's experience are inextricable from the Turkish context, the overarching themes in the process of her journey to insurgency– including the relationship between conflict and economic opportunity, social and familial networks, and insurgent group structure– are of fundamental importance to our understanding of conflicts around the world.

The answers to these questions speak to a vast and ever-growing literature that seeks to understand the factors that drive people to take up arms and the policies that may persuade them to lay them down. However, these efforts have been stymied by the fact that conflict, by nature, is exceptionally difficult to study. Other facets of the human condition can readily be analyzed through interviews, randomized control trials, ethnographies, and regularly collected censuses. But because none of these approaches are feasible in a war zone, there is a dearth of detailed empirical evidence on the conduct of conflict. Even some of the most widely available statistics such as Gross Domestic Product are systematically missing for countries experiencing conflict (Gleditsch, 2002).

At the same time, one of the defining characteristics of the 21st century has been the exponential growth of information generated by individuals and organizations. There have been academic efforts to harness these vast quantities of data under the banner of Computational Social Science, "an interdisciplinary field that advances theories of human behavior by applying computational techniques to large datasets from social media sites, the Internet, or other digitized archives such as administrative records." (Edelmann et al., 2020: 67) But this field is still in its infancy, and academic research has largely failed to keep up with this torrent of information despite the fact that buried within it lies evidence that can contribute novel insights towards longstanding debates.

A central contribution of this thesis is the development of computational methods that are able to harness these large quantities of data and overcome many of the key challenges of conducting empirical research on armed groups. Though interviewing in a warzone is–for obvious reasons– difficult, the life history of Gülnaz Ekinci was constructed without leaving Oxford by triangulating open source material available online. While randomized control trials cannot be conducted in between battles, merging insurgent obituaries with a leaked population registry allows for direct comparisons between militants and random samples of ordinary citizens. Census enumerators are unlikely to gather meaningful data from the battlefield, but imagery from earth observing satellites can provide measurements of agricultural and economic activity spanning decades at an extremely high spatial and temporal resolution, regardless of the bloodshed below. Unlike firms, militant groups don't often publish organograms of their structure– indeed, they often go to great lengths to conceal this information. Yet using nothing but a large trove of photographs published online by the PKK, the organizational structure of the group can be reconstituted.

In what follows, I develop computational methods to gather, process, and analyze data on three key facets of the conflict between the PKK and the Turkish government. The remit of these studies is encapsulated by three broad questions: What factors lead people to join the PKK? How have policies aimed at reducing recruitment fared? And why have military means failed? These chapters yield new insights into the nearly 40 year old war between the Turkish government and the PKK, and though many of the empirical results are contextually contingent, they speak to much broader debates in the study of civil conflict. Following a brief contextual note on the Kurdish conflcit, each chapter is summarized in greater detail below.

The Kurdish Conflict

Following the establishment of the Turkish Republic in 1923, the government undertook an effort to create a nation-state modeled after European powers such as France (Park, 2013). This involved the writing of a constitution, a strong emphasis on secularism, and most importantly a policy of "Turkification", whereby "Turkish ethnic identity has been strictly imposed as a hegemonic identity in every sphere of social life" (Aktar in Ulker, 2005: 29). Later that year, a group of Kurdish tribes led by Sheikh Said rebelled against the Republican government with the goal of establishing an "independent Kurdistan" (Olson, 2000: 69). The rebellion lasted nearly four months, involved the siege of the region's largest city, Diyarbakir, and led to 8,000 casualties before Sheikh Said was defeated (Ibid).

For much of the 20th century, Southeastern Turkey saw sporadic bouts of armed insurrection by the Kurds, followed by devastating military campaigns by the government. The development of a unified Kurdish resistance movement came in 1974, with the establishment of the Kurdistan Worker's Party (Patiya Karkerên Kurdistanê, henceforth PKK) by Abdullah Öcalan (Özcan, 2013). Thereafter, the cycles of revolt and repression were replaced with a low-intensity conflict beginning in 1984, peaking in the 1990s, and continuing to this day. Following the breakdown of a 2013-2015 truce, hostilities between the PKK and the Turkish government have intensified almost to the level of the conflict's peak in 1994. Olson (2013: 32) notes that "more Turkish military actions have been carried out in Kurdistan than in any other area of Turkish concern, foreign or domestic".

Chapter 1: Youth, Urbanization, and Insurgent Recruitment in Turkey.

Gülnaz Ekinci joined the PKK at a young age, and was born in a rapidly growing city to parents who had recently emigrated from a rural area. On the eve of her death, the city she was born in had more than doubled in size, and the house she grew up in had been bulldozed to make way for a new development.

Indeed, the notion that large urban youth cohorts are a key driver of civil conflict is widely held among policymakers and scholars alike, particularly since the youth-led protests in cities across the middle east evolved into the Arab Spring (Hvistendahl, 2011). Yet, in a review of the literature associating urban "youth bulges" with conflict, Cramer (2011: 2) concludes that "we still know too little empirically" to make any robust conclusions thereon. This problem is rooted in the inaccessibility of data, as Humphreys and Weinstein (2008) argue that in order "to properly assess competing explanations, we need a research design that permits a comparison of the characteristics of participants and nonparticipants." Though such data is exceedingly rare in the context of an active conflict, by merging thousands of PKK obituaries (like Gülnaz's) with a leaked population registry, this study is able to examine spatial and temporal trends in the recruitment of urban youth at the level of individuals. To my knowledge, this study represents the first ever large-scale quantitative comparison of insurgents to both random and demographically matched samples of the entire adult population of a country. Thus, the first chapter of this thesis leverages an unprecedentedly powerful research design to conduct a detailed exploration of the factors associated with urban youth recruitment to the PKK.

This chapter fills two gaps in the literature connecting large urban youth cohorts and conflict. The first is that the literature on "youth bulges" curiously neglects many factors that are widely recognized in economics and sociology as affecting key outcomes for young people. These include social and familial variables such as peer effects, birth order, and family size. Original findings from this study suggest that second born children and children from larger families are more likely to join the PKK, which aligns with econometric studies of youth crime. Dense spatial clustering of PKK recruits within urban neighbourhoods compared to the general population is suggestive of peer effects in the process or recruitment.

The second gap filled by this chapter involves bringing new empirical evidence to a fraught debate on the relationship between urbanization and conflict, as existing studies have yielded highly mixed results. Most studies simply observe whether or not there are more violent clashes in countries where there are growing cities, which does little to further our understanding of the relationship between urbanization and civil conflict. The use of cities or grid-cells as units of observation does not address the need to analyze the causes of migration, the nature of urbanization, and the mechanisms through which these may lead to civil strife. I address this gap in the literature through a detailed comparison of how militants' migratory patterns and spatial clustering differ from those of the general population. Findings align with previous cross-country studies on the relationship between displacement and conflict (Salehyan and Gledistch, 2006). Having explored many of the social, familial, and political factors associated with the recruitment of urban youth to the PKK, the next chapter focuses on the effects of a large agricultural development program on Kurdish separatism.

Chapter 2: Assessing the impact of the Southeastern Anatolia Project on Kurdish Separatism.

Eight kilometers from the arid farming village that Gülnaz's parents emigrated from, the Turkish government built one of the largest irrigation dams in the world. Though this dam was only completed in 2017, it is one of 22 that the government has built in the region over the past four decades as part of the Southeastern Anatolia Project (Guneydogu Anadolu Projesi, henceforth GAP), a regional development program that seeks to irrigate an area around half the size of Lebanon (GAP, 2016).

The stated purpose of the project was to "improve the level of income and life quality of the local population" (GAP, 2018). However, there was also a hope that "GAP could turn Kurds into Turks" and that "GAP will help to decrease, if not eliminate, the appeal of the PKK" (Jongerden, 2010: 137; Olson, 1996: 96). A leaked U.S. State Department cable indicates that Turkish policymakers (including then-President Abdullah Gul) believed that "If Kurds are gainfully employed, have better educational opportunities, and see increased levels of infrastructure development throughout their region, their affinity for the terrorist PKK will wane further. A top priority is ensuring completion of the massive Southeastern Anatolia Project (GAP) within five years." (Wikileaks, 2008b). Ethnographic studies conducted in the largest irrigation scheme created by the project found a "heightened sense of state legitimacy as a function of irrigation access"; in the words of one farmer, "our view of the state changed positively... we had hatred before, but now they started investing in the Southeast" (Harris, 2016: 9). Though many have hypothesized that GAP would improve rural Kurds' trust in the Turkish state enough to dissuade them from joining the PKK, this claim has yet to be empirically investigated. Thus, the purpose of this chapter is to answer the following question: how has irrigation wrought by the Southeastern Anatolia Project impacted Kurdish separatism?

To answer this question, I create 5km and 10km gridded datasets on irrigation and Kurdish separatism from 1985-2019, exploiting exogenous topographical variation in the distribution of irrigation schemes. Consistent results were derived from both cross-sectional Instrumental Variables approaches and spatial panel models, and were robust to the inclusion of a wide array of highly detailed political- economic control variables, alternative measures of irrigation and conflict, and fixed effects. I find that between 1985 and 2019, a 27 km2 increase in irrigated area decreased the likelihood of experiencing a conflict event in a 100 km2 grid-cell in a given year decreased by 49% relative to the mean. A district-level analysis suggests the decoupling of agricultural income from rainfall as a likely mechanism. Clashes are more frequent following a poor harvest, and yields for all major crops except irrigated cotton are highly sensitive to rainfall. However, there is substantial treatment heterogeneity related to land inequality.

These results make an important contribution to the literature on the microeconomics of violent conflict through a detailed investigation of the political economy of Kurdish separatism in Southeastern Turkey. The most highly cited paper on civil war proposes that an individual's decision to join a rebellion is a function of "whether economic opportunities are so poor that the life of a rebel is attractive to 500 or 2,000 young men" (Fearon and Laitin, 2003: 88). Yet this mechanism remains poorly understood and largely untested. Though Cederman, Weidmann, and Gleditsch find cross-national evidence that horizontal inequalities between ethnic groups increase conflict likelihood, micro-level analyses thereof are still scant. By examining heterogeneity in the relationship between rainfall and agricultural income introduced by a topographically determined irrigation project, I demonstrate the potential of development projects to reduce violent conflict; a finding with implications that reach far beyond Turkey.

Though the first two chapters provide detailed insights into the factors associated with insurgent recruitment, the use of obituaries means that the sample is likely biased towards lower-ranking individuals within the PKK as they are killed more frequently and in greater numbers than elites. However, the PKK is not simply an agglomeration of foot soldiers, but a complex hierarchical organization. As such, the final chapter turns its attention to an oft-neglected dimension of most conflicts: the structure of the insurgent group itself.

Chapter 3: Rebellion as Complex Network: Social Ties and Hierarchy in the PKK

By the time she was killed, Gülnaz had risen to the rank of Commander and the loss of her experience and leadership was doubtlessly a blow to the PKK. In the years since her death, the Turkish government embarked on a policy of targeting senior leadership through drone strikes and raids. Yet the PKK has not only managed to survive, but match this escalation. The organizational structure of a militant group governs many of its fundamental characteristics including its robustness to different counterinsurgency strategies, willingness to pursue negotiated settlements, and factionalization. There is a growing recognition that the outcomes of conflicts cannot be abstracted from structural characteristics of the insurgent groups that are involved.

The primary contribution of this chapter is a novel approach to the study of insurgent group structure, which is motivated by three principles. First, empirical assessments of insurgent group structure are primarily impeded by a lack of data. Second, that one source of data that is particularly abundant– often generated by militants themselves– are photographs published on the internet. Third, that co-appearance in photographs is recognized as indicating some level of meaningful social interaction, as they have been used to construct social networks in other settings. The intuition behind this approach can be summarized as follows: If two people are co-present in a photograph, the nature and social significance of the link between them is unknown. However, if the same person systematically appears with others in a large number of photographs, there is probably an overarching reason. In this paper, I ask: To what extent can co-appearance in militant photographs be used to understand the social, political, and organizational forces that structure an insurgent group?

I answer this question through an original collection and analysis of nearly 20,000 images scraped from the PKK's obituary website and a novel image processing pipeline. I extract facial encodings from these photos, use unsupervised clustering to group images of the same individual across multiple images, and create a social network graph based on co-appearance. In three analytical sections, I explore how various properties of the PKK Image Co-Appearance Network (PICAN) align with external information about the PKK. Section 5 shows that a qualitative analysis of the nature of an individual's co-appearances can yield information on functional and factional divisions within a rebel group. Section 6 demonstrates a consistent relationship between an individual's rank and their centrality in the co-appearance network. Section 7 shows that the co-appearance network may even approximate the general structural characteristics of the insurgent group via a close reading of the topology of the PICAN in reference to qualitative work on the PKK's history. The substantive findings suggest that the PKK is not hierarchical enough to be incapacitated by the targeting of a handful of top officials, due to a large number of mid-level officers such as Gülnaz. Indeed, conflicts with structurally similar groups such as the FARC in Colombia have only been resolved through negotiated settlement. Thus, this study concludes that the conflict with the PKK is unlikely to be conclusively settled through military means.

Foreword

Few human activities are both as destructive and opaque as civil war. Yet developing a deep understanding of the dynamics and causes of these conflicts is a prerequisite for their resolution. Academic efforts to achieve this understanding have made significant progress, but access to the information necessary to do so is hampered by the fundamentally chaotic nature of war. Blattman and Miguel (2010: 8) contend that "the most promising avenue for new empirical research is on the subnational scale, analyzing conflict causes, conduct, and consequences at the level of armed groups, communities, and individuals." The ultimate goal of my research is to leverage technological advances that enable the collection and analysis of vast quantities of data in conflict zones that was once impossible. These three chapters utilize large quantities of fine-grained data and develop cutting-edge analytical techniques to yield novel insights on the conflict between the Turkish government and the PKK. Together, they provide new empirical evidence that speak to ongoing debates in the literature on civil wars, and set out novel approaches in computational conflict research.

Chapter 1

Youth, Urbanization, and Insurgent Recruitment in Turkey

1.1 Introduction

One of the most pervasive understandings of civil unrest attributes violent conflict to the emergence of large urban youth cohorts. Hvistendahl (2011) notes that a relative consensus emerged among western policymakers that the Arab Spring could be attributed largely to these demographic changes. The scholarly endeavor to understand the causes and conduct of civil conflict has been dominated by cross-country regressions, and while these studies may reveal general trends and relationships between certain variables and the onset of civil conflict, they cannot yield detailed insights into the nature of these relationships. Though research on conflict is becoming more detailed and is increasingly being conducted on the level of sub-national geographic units, there are very few quantitative studies of the factors associated with participation in an insurgency conducted at the level of individuals (Verwimp et., al. 2019). Yet Humphreys and Weinstein (2008) note that in the study of civil conflict, in order "to properly assess competing explanations, we need a research design that permits a comparison of the characteristics of participants and nonparticipants."

This paper leverages an unprecedentedly detailed dataset to conduct individual-level comparisons between insurgents and the general population to study the factors associated with the recruitment of urban youth in Turkey. I merge insurgent obituaries with a leaked population registry to create a novel anonymized dataset containing a rich array of individual-level demographic variables for deceased militants. These include the precise locations of birth and residence, household size, number of siblings, as well as village- and neighbourhood-level population. Individual-level data are geolocated based on their birthplace and registered address, and merged with data on ballot-box level election outcomes, satellite-derived nighttime lights, and conflict incidence. The high degree of detail in this dataset enables novel empirical insights, including individual-level comparisons between insurgents and random samples of the Turkish population, as well as the creation of a matched sample of non-recruits who share the same demographic characteristics as recruits. The overarching question guiding this analysis is: what factors are associated with the recruitment of urban youth to the PKK?

This question is answered in the following four sections. Section 2 identifies gaps in the literature on the relationship between urbanization, youth, and conflict. Section 3 details the construction of the insurgent recruitment dataset, as well as potential concerns related to non-random selection and research ethics. Section 4 summarizes the collection of other covariates and discusses summary statistics. Section 5 presents the results of logistic regression models comparing recruits to non-recruits, and discusses each result in greater detail. Five sub-sections analyze the relationship between urbanization, youth, and insurgent recruitment at progressively finer scales, guided by a number of latent questions in the literature. Section 5.1 explores country-level trends in internal migration, finding evidence of conflict-induced displacement as a driver of both urbanization and militancy. Section 5.2 examines city-level characteristics of migrants' destinations and their relationship with insurgent recruitment, finding no support for the proposition in the literature that growing peri-urban areas are disproportionately conflict-prone. Section 5.3 reveals dense spatial clustering at the neighbourhood-level among militants in urban areas, likely indicative of peer effects in the process of insurgent recruitment. Household-level factors including family size and birth order are analyzed in Section 5.4, and are found to play an important role in the recruitment of young people. Section 5.5 interrogates the dismissal of individual-level political grievances in the study of insurgency, finding significant relationships between political engagement and militancy, as well as a spike in recruitment tied to a pivotal development in Kurdish politics. Together, these findings provide novel insights on a range of factors often hypothesized to affect the recruitment of urban youth into conflict, but that are quantitatively understudied.

1.2 Youth, Urbanization, and Conflict

While research into the economic drivers of conflict has proliferated (Verwimp et. al., 2019), there are two key dimensions of insurgent recruitment that are difficult to operationalize quantitatively and consequently empirically understudied. The first involves social and familial factors such as birth order, family size, and peer effects, all of which are well studied in economics and recognized as affecting a wide range of outcomes among young people. The second involves the causal mechanisms linking urbanization and conflict. In the remainder of this section, I explore the gaps in these literatures and highlight how the present approach–geolocated individual-level comparisons between militants and the general population– can address them.

1.2.1 Social and Familial Factors

"Youth bulge" theory posits a causal relationship between the proportion of young men in a society and the risk of violence (Urdal, 2006, 2008; Cincotta, Engelman, et. al. 2003; Goldstone, 2002). This belief is enshrined in much of the seminal theoretical work on violent conflict, including in Fearon and Laitin (2003: 86), who argue that "young males have physical and perhaps psychological characteristics that make them apt guerrillas". Similarly, Collier, Hoeffler, and Roehner (2008: 22) contend that "the proportion of young men in the society is a good proxy for the proportion of the population psychologically predisposed to violence and best-suited for rebel recruitment".

Despite this emphasis on youth, the theoretical frameworks designed to explain why individuals join insurgencies have been implicitly developed for adults: opportunity cost mechanisms presuppose that an individual is in the labour market, and grievance-based explanations assume a degree of political consciousness. Though teenagers are not precluded from being politically conscious or labour market participants, insufficient attention has been paid to how these pressures might differentially affect younger people. Furthermore, factors that are important to childhood development but less relevant for adults– family size, birth order, social dislocation– are largely absent in the theoretical and empirical literature on insurgent recruitment.

Many economic pressures associated with insurgent recruitment are likely mediated through familial dynamics. Parents' investment in (and therefore a young person's access to) education is likely a function of the number of children in the family, and children face different pressures to participate in the labour market depending on their birth order and role in the family. A recent and influential study by Doyle et. al. (2020: 97) found that "involvement with the juvenile justice system is found to be on the order of 30%–40% higher compared with the mean level of involvement among firstborn boys in both Denmark and Florida. For example, 7.2% of firstborn boys in Denmark are sentenced to prison by the age of 21, and we estimate that secondborn boys have a rate that is 2.4 percentage points higher: a 33% difference when comparing brothers in the same household. These effects are particularly strong among more severe violent crimes (36%)." Thus, adjacent literature in economics on the effect of birth order on labour market outcomes and violent crime suggests that these factors could be highly relevant to insurgent recruitment.

Similarly, there has been substantial econometric research into the influence of peer effects among young people on a range of outcomes including test scores (Burke and Sass, 2011; Zimmerman, 2003), the likelihood of joining social groups such as fraternities (Sacerdote, 2001), and even adolescent overweight (Trogdon et. al. 2008). Qualitative studies of youth participation in conflict affirm the importance of social networks in the process of recruitment. In interviews with former youth combatants in the DRC, Pakistan, Afghanistan, Northern Ireland, and Colombia, Brett and Specht (2004) found that one of the most common themes among young recruits across contexts was having a friend or family member who was already a member of the rebel group in question; they even highlight multiple instances in which groups of friends will join together. Though Edgerton (2022) studies the effect of kin and peer ties on suicide bomber mobilization among Islamic State militants, quantitative investigations of peer effects in the process of insurgent recruitment are lacking. Thus, there is a significant gap in the quantitative literature on youth participation in armed conflict.

Qualitative evidence reaffirms the importance of social and familial factors that are of particular relevance to young people. However, because these factors are hard to measure– indeed, they require a large number of individual-level observations– they are understudied. Data compiled for the present study offer unprecedented quantitative insights into the recruitment of young people into armed conflict, including the effect of familial variables such as the number of siblings, their birth order, and relative ages, as well as potential peer effects.

1.2.2 Urbanization

Claims regarding the volatility of large youth cohorts are often linked to another significant demographic shift in the developing world: rapid urbanization. Juarez and Urdal (2020) note that "there is an increasing concern that rapid urban population growth around the globe could lead to increasing levels of criminal as well as political violence". Similarly, in a paper titled "Rapid Urbanization and the growing threat of Violence and Conflict: a 21st Century Crisis", Patel and Burkle (2011: 194) assert that "as urban slums become a haven for criminal elements, youth gangs, and the arms trade, they also create insecurity for much of the population." Thus, as large population movements lead to overcrowding and poor living conditions in cities, conflict is taken to be a natural consequence (Goldstone, 2002). This relationship, however, remains more assumed than proven.

Empirical analyses of the relationship between urbanization and conflict are rare, and findings are often contradictory. A country-level analysis by Esty et. al. (1998) suggested that the risk of conflict nearly doubled in countries with above-average levels of urbanization but below-average GDP. A city-level study by Buhaug and Urdal (2013), however, found no association between urban growth rates and the incidence of violent conflict. Clashes were instead associated with economic shocks, ongoing civil conflict, and a lack of consistent political institutions. However, a recent study by Gizelis, Pickering et. al. (2021) found increased conflict incidence in peri-urban areas using sub-national grid-cells. All three studies analyze the African continent over roughly the same time period. All three present hypotheses about the reasons behind the observed trend of increasing urbanization, but none are able to measure them directly. Though the trend towards spatial disaggregation in these studies addresses some of the shortcomings of earlier cross-country work, it is subject to new, and perhaps more serious conceptual concerns that may help explain these contradictory findings.

The first is the use of conflict incidence as a dependent variable with which to study the relationship between urbanization and conflict. There is likely to be a significant disconnect between the location of a conflict event and the geographies that are relevant to its perpetrators, particularly in urban areas. If there is an explosion in the financial district of a capital city, are the social, economic, and demographic characteristics of the immediate vicinity relevant to our understanding of the reasons the bomb was planted? Perhaps, if the assumption is that the perpetrator was born and raised nearby. Yet this is not an assumption that can be made– by looking simply at the location of this conflict event, we do not know whether the bomb was planted by a university student born and raised nearby or a farmer who recently migrated due to drought. Each of these scenarios would radically change our understanding of the drivers of conflict, and neither the relative proportions of these demographic groups at the national level nor the exact location of the conflict event further our ability to determine which scenario is the most likely.

This relates to the second conceptual flaw in these studies, which is the use of simple measures of urban growth as an independent variable. The decision to move to a particular city is itself the product of a number of factors that are directly relevant to insurgent recruitment: the pursuit of economic opportunities, forced displacement, weather shocks, among others. As such, the characteristics of an individual's birthplace, the city– even the neighbourhood– of destination, and the differences between them are of critical importance. Furthermore, these variables likely interact in significant ways with many of the familial and social factors identified above; because individuals tend to migrate to areas in which they have existing diasporic or kin ties (Massey, 1990), the social transmission of ideas is likely to be spatially mediated. Certain neighbourhoods may comprise socially and familially embedded individuals who migrated from similar areas for similar reasons. Thus, peer effects would manifest themselves in the form of spatial clustering of insurgent recruitment at the local level. Ultimately, the nature of urbanization is far more relevant to civil unrest than the quantity thereof; while the existing literature focuses overwhelmingly on the latter, this paper examines the former.

Simply observing whether or not there are more violent clashes in countries where there are growing cities does little to further our understanding of the relationship between urbanization and civil conflict. The use of cities or grid-cells as units of observation does not address the need to analyze the causes of migration, the nature of urbanization, and the mechanisms through which these may lead to civil strife. Indeed, even the rare individual-level analyses of insurgent recruitment do not necessarily fill this gap. Humphreys and Weinstein (2008: 437) note that these studies often select on the dependent variable by only interviewing current or former combatants, yet "to properly assess competing explanations, we need a research design that permits a comparison of the characteristics of participants and nonparticipants." I address this gap in the literature through a detailed comparison of how militants' migratory patterns and spatial clustering differ from those of the general population using both matched and random samples drawn from the entire adult population of Turkey.

1.3 The Turkish Context

The Kurdish insurgency in Turkey provides an appropriate setting in which to study the relationship between urban youth cohorts and insurgent recruitment. At 76%, the rate of urbanization in Turkey is substantially higher than that of many of the Sub Saharan African countries traditionally included in studies of urban growth and conflict (World Bank, 2020). Furthermore, classified U.S. State Department documents published on Wikileaks reveal "youth bulge" theory to be one of the primary lenses through which the Kurdish insurgency has been understood (Wikileaks, 2010). A cable addressed to the U.S. Central Intelligence Agency transcribes a meeting between political officers of the U.S. Consulate in Istanbul and Kurdish intellectuals to discuss the radicalization of Kurdish youth. The cable posits that "most of the radicalization is among Kurdish youth who have no particular political goal and are frustrated with what they view as a hopeless situation", and that "The youth bulge suggests that radicalization will continue". Quoting one of the Kurdish intellectuals, the cable concludes that "this population is 'like a treasury' of recruits for the PKK."

1.4 Methodology

This section details the methodology behind the construction of an unprecedentedly detailed dataset of insurgent recruitment, discusses the ethical implications thereof, as well as potential sources of selection bias. The recruitment dataset is compiled in three main steps: first, by scraping the PKK's obituary website for basic biographical information. Second, by matching obituaries to records in a leaked copy of the Turkish Citizenship Registry (Merkezi Nüfus İdaresi Sistemi, or MERNIS) database. Third, by geocoding address information to enable geospatial analysis. All of the data and methods used herein are open source.

When a fighter is killed, an obituary is posted on a website run by the PKK's military wing, the HPG. Under a banner reading "Our Martyrs are our Glory", the short obituary contains pictures and basic information regarding the fallen fighter including their full name, parents' names, as well as the dates and places of birth, recruitment, and death. Not all obituaries contain all of these fields; the PKK stopped reporting birth and recruitment information after 2012, and other fields are sometimes blank. The first step in this analysis is to scrape the biographical data contained in these obituaries, creating a database containing biographical information on 3,664 PKK militants who were killed in action.

While the PKK's online obituaries contain basic information for each recruit, the amount of detail can be significantly augmented using a leaked copy of the Turkish citizenship registry. A SQL dump file of the registry was uploaded to the internet in 2016, containing information on 49 million individuals who were adults in 2010. Using the basic biographical information provided in the PKK obituaries, individuals can be matched to records in the citizenship registry. Fuzzy matching is used for names, accounting for possible variations in spelling. A weighted average of these individual field matches is computed to yield a score for the overall match. A match is considered genuine if the obituary and the record have the same first name, last name, parents' names, and birth date, allowing for minor variations in spelling and small discrepancies in the birth date. To enable geospatial analysis, addresses are geocoded using the Google Maps API. In the interest of privacy, coordinates are offset randomly by up to 1km.

1.4.1 Selection Bias

Though this procedure yields a highly detailed dataset on insurgent recruitment, selection bias is a concern. There are three sites at which potentially non-random selection occurs in the generation of the final dataset: inclusion in the PKK obituary website, overlap with the MERNIS citizenship registry, and the process of record linkage between the two.

The use of obituaries to construct the initial sample of PKK members introduces selection bias. In a study that also makes use of PKK obituaries, Tezcur (2016) hypothesizes that two main groups of insurgents are at risk of being excluded from the sample. The first group is active militants, whose exclusion could bias the sample towards the overrepresentation of recruits from lower socio-economic backgrounds who were more often sent on dangerous missions than the more educated recruits. The second group is militants who were killed but whose information was inaccessible or unpublished. The decision to not publish certain obituaries could result from the fact that a recruit was executed by the PKK, a fate that befell up to 900 insurgents by certain estimates (Yilmaz, 2014: 192). Tezcur (2016: 192) hypothesizes that this would primarily affect educated recruits who might be more likely to question the PKK's top-down structure. Finally, there are some discrepancies between the obituary data and the citizenship registry data among clear matches. One instance involves the doctoring of a birth date by one year to mask the fact that an individual was recruited as a child. Overall, these discrepancies are relatively rare and have a median of zero; while some manipulation of the data by the PKK is certainly possible, it does not appear systematic. Nevertheless, subsequent analysis of matched individuals uses MERNIS-derived values. As such, this type of manipulation can only lead to lower match rates rather than biased variables.

A second source of non-random selection involves the temporal overlap between the records in the MERNIS citizenship registry database and the obituaries. The MERNIS database contains 48 million records, corresponding to roughly the entire adult population of Turkey in 2010. As such, the database does not include anyone born after 1992, introducing non-random missingness due to the failed matching of obituaries for individuals born after this year. This problem becomes increasingly severe with time, as more recent obituaries are more likely to contain individuals born after 1992. Though the PKK stopped reporting birth dates in 2012, 950 records have both the year of birth and year of death. The difference between these values yields a distribution d ($\mu = 26$, $\sigma = 5.1$) of the ages at which militants are killed. Assuming this distribution is relatively constant over time, the estimated proportion of obituaries that cannot be matched due to temporal incongruence between the datasets can be expressed as follows:

$$m = \frac{\sum_{i=2001}^{2019} P(d < i - 1992) \times n_i}{n} \tag{1.1}$$

Where d is the age-at-death distribution, i is a year between 2001 and 2019 (the range of years covered by the obituaries), and n is the number of obituary records. For a given year, the estimated number of missing records is the proportion of the age-at-death distribution that is smaller than the number of years elapsed since 1992, multiplied by the number of obituaries posted that year. The sum of this value across all years yields the estimated number of missing records (893), suggesting that 25.2% of PKK obituaries cannot be matched with the MERNIS database. A further 17.7% of the obituaries were posted before 2010, and could not be matched as deceased individuals are removed from the MERNIS database. In total, 43% of obituaries could not be matched due to the timespan covered by the citizenship registry data. The matched obituaries are thus less representative of both older and more recent trends in PKK recruitment. Subsequent analysis of this sample will accordingly focus primarily on cross-sectional trends.

The remaining 30% of unmatched obituaries likely result from stringency in the matching process. Priority was placed on having a smaller sample of high-certainty matches, rather than having a larger sample that included false positives. As such, only records with multiple exact matches or very close matches in terms of name, date of birth, and parents' names were used. In the process, many near but ultimately indeterminate matches are excluded because false positives are more dangerous in this context than false negatives. Despite shrinking the sample size, missingness resulting from discrepant spellings is likely to be largely random, and is offset by the creation of a more precise sample. Lowering the matching threshold would increase the raw number of matched obituaries, but introduce noise.

As such, the resulting dataset should be understood as a sample of PKK recruits that is likely biased towards lower-ranking individuals within the organization who were born prior to 1992, and died after 2010. The resulting limitations are that the dataset cannot deliver precise insights into the nature of more recent recruitment patterns, and that inferences drawn from this sample may not reflect the experiences of higher ranking individuals.

1.4.2 Research Ethics

Traditional research into insurgent recruitment typically employs in-person interviews, which carry significant ethical and practical costs to both the researcher and the participant. Researchers are often required to travel to extremely dangerous areas, thereby putting themselves in harm's way. Because participants are alive and have often either left the insurgency (and thus may be targeted for retribution) or have been captured, there are serious ethical implications of conducting in-person research on insurgent recruitment. As a result, detailed research thereon is severely lacking. Rather than creating new sensitive data (e.g. by interviewing), I only leverage public domain data that anyone with an internet connection can already access. Furthermore, by limiting the study to deceased individuals, it is unfeasible to cause further harm to participants. Thus, the current approach entails fewer risks to both researchers and participants than the status quo of interview-based research into individual-level participation in insurgency.

Though the MERNIS dataset has previously been used in academic research, including a study by Rocher, Hendrickxs, et.al., (2019) published in Nature, I take additional steps to reduce the sensitivity of the dataset via a nine-step anonymization process. This process generates a number of encrypted relational tables that do not individually contain enough information to be considered sensitive; "NAME_DB" above contains only a list of names, genders, and dates of birth. "GEO_DB" contains a list of anonymized addresses. Database user privileges are configured such that each table requires a separate password to be accessed. Ethics approval for the present research following the above methodology was granted by the Oxford Central University Research Ethics Committee (CUREC), (ref: ODID C1a 21-064). The full text of the CUREC application and the approval are included as supplements, and a diagram of the anonymisation workflow is available in the appendix.

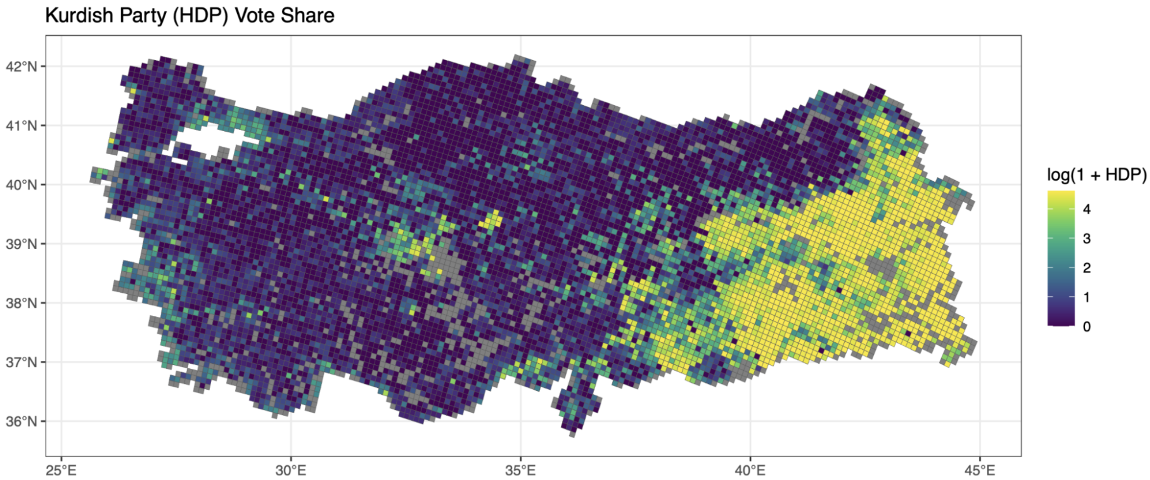
1.4.3 Data Collection

This section details the collection of a range of social, political, and demographic control variables that form the basis of the subsequent quantitative analysis. To properly address the political dimension of insurgent recruitment, I collect millions of ballot-box level election results in which the main Kurdish party (*Halkların Demokratik Partisi*, henceforth HDP) participated. To test claims in the literature that peri-urban areas are particularly conflict-prone (Gizelis, Pickering et. al., 2021), I create a longitudinal measure of urban growth using nighttime lights data. The inclusion of fine-grained MERNIS-derived population counts enables an assessment of the number of PKK recruits per capita down to the neighbourhood level. Finally, to better understand the drivers of internal migration and urbanization– in particular conflict-induced displacement– I rely on conflict incidence data. Following the description of the data collection process for these variables, I provide summary statistics.

1.4.4 Ballot-Box Level Political Data

Though there is considerable disagreement in the literature regarding the significance of political motivations for rebellion (Collier and Hoeffler, 2004), interviews with Kurdish militants often explicitly mention party-political engagement as an important factor in their decision to participate in armed struggle (Aytekin, 2019). Recently, the arbitrary arrest and detention of Kurdish HDP politicians has become a significant source of grievances for both the PKK and wider Kurdish community in Turkey (Reuters, 2021).

Figure 1.1: Kurdish Party (HDP) Vote Share



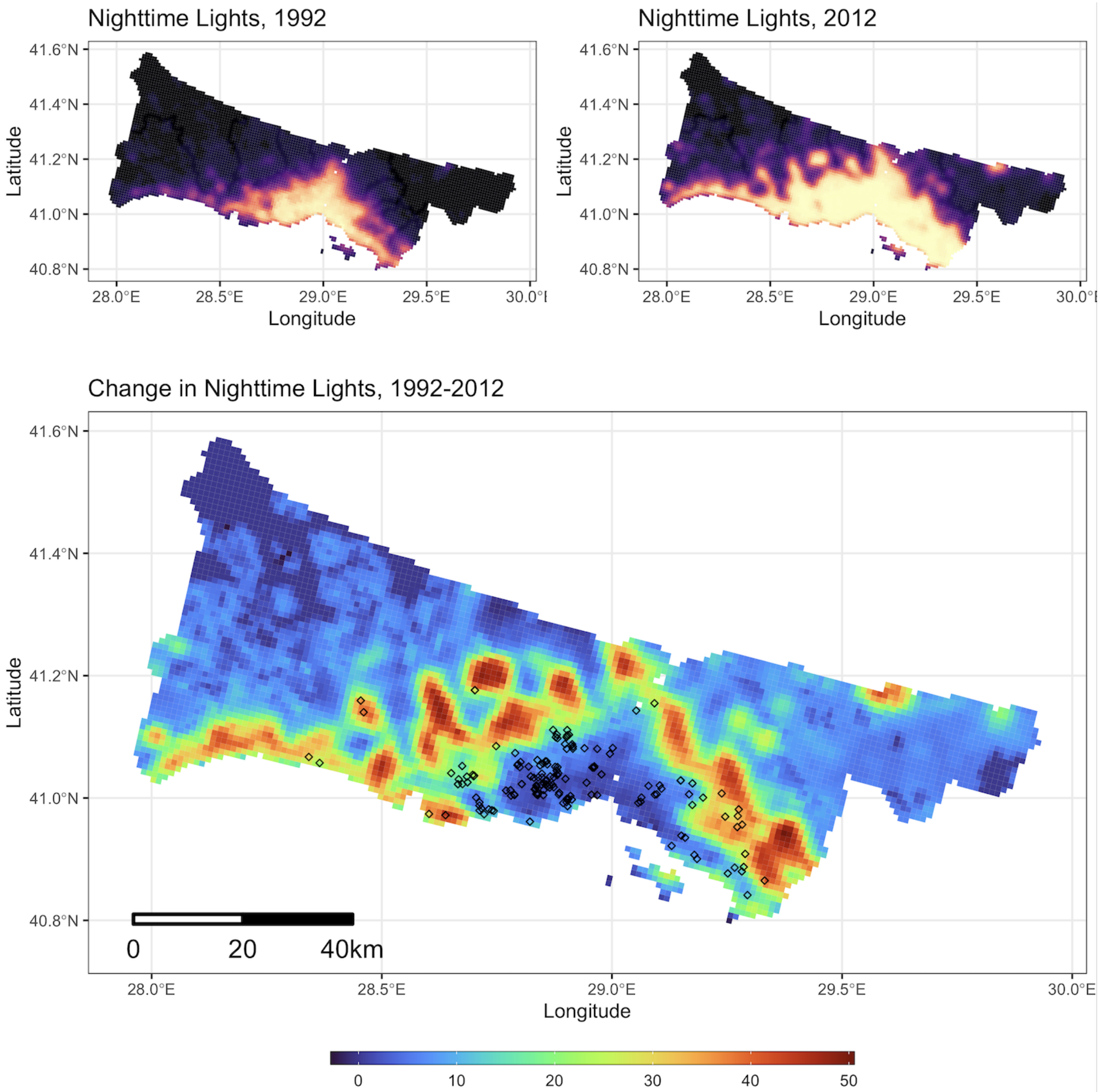
This map shows the log of the mean vote share cast for the Halkan Demokratik Partisi (HDP) in the 2015 General Election at the level of 10km by 10km grid cells. The HDP is the main Kurdish political party in Turkey.

2.9 million ballot-box level election results were scraped from the Turkish Supreme Elections Board website, and geocoded using the Google Maps API. Results from the 2015 Parliamentary Election were isolated due to the salience of the Kurdish issue therein. The elections occurred in the context of a peace process between the Turkish government and the PKK, leading to a surge in popular support for the HDP and marking the first time the party met the 10% parliamentary threshold (Baser and Ozerdem, 2019). The main political variable used in subsequent analysis represents the 2015 vote share of the HDP, shown in Figure 1.1.

1.4.5 Nighttime Lights

Nighttime lights data are used to provide information on urbanization and economic development at a high spatial resolution. Nighttime lights data were collected from the Defense Meteorological Satellite Program (DMSP), and reflect the average radiance displayed by a cell in a given year. Though a number of existing measures of urban area exist, these are all time invariant. To generate a measure of new urban growth, the difference in nighttime lights between 1992 and 2012 is measured. Figure 1.2

Figure 1.2: Urban Growth, Istanbul



The top two panels show the growth in Nighttime Lights radiance in Istanbul between 1992 (left) and 2012 (right) at the level of 1km^2 grid cells. The bottom panel visualizes the difference between the two images, highlighting new urban growth during this period. Black circles represent PKK recruits.

demonstrates this process being applied to Istanbul. The image on the top left shows Istanbul's nighttime radiance in 1992, while the image on the top right shows radiance in 2012. The image on the bottom shows the difference between these images, highlighting a ring of peri-urban growth surrounding the city. An important consideration is that this process highlights areas of new urban growth; Istanbul's city centre, which was luminous both in 1992 and 2012, registers a low value on the nightlights growth variable. Similarly, forested areas to the Northwest of the city were dark in 1992 and dark in 2012, and therefore also present low values for nighttime lights growth.

1.4.6 MERNIS-derived variables

A number of control variables were constructed using the entirety of the MERNIS database. Detailed population measures were created for the city, neighbourhood, and village level by counting the number of records in each locale. Because more populous areas are likely to have larger numbers of recruits, fine-grained population counts can highlight areas in which there is a disproportionate number of recruits per capita. In addition to these, a measure of household size was created for each individual by counting the number of individuals at a given address with the same last name. Household size has long been recognized as an important component of family structure linked to key outcomes for young people such as educational achievement and delinquency (Nye, 1958; Demuth and Brown, 2004).

1.4.7 Conflict Incidence

Though the causes of urbanization are myriad, conflict-induced displacement has been a one of the main causes of internal migration in Turkey (Çelik, 2013). To examine the relationship between urbanization and conflict, georeferenced historical data on conflict incidence were collected from the Uppsala Conflict Data Program (UCDP). The primary variable generated using these data is a measure of proximity to conflict events. This was generated by calculating the mean inverse distance between an individual's place of birth and all conflict events prior to 2012.

1.4.8 Summary Statistics

Table 1.1 shows summary statistics including the mean and standard deviation for each of the variables described above. Columns disaggregate the statistics by sample. The "PKK Members" sample is composed of obituaries that were successfully matched with records in the MERNIS database. The "Random Sample" is a simple random sample of records from the MERNIS database. The "Siblings" sample comprises the

Table 1.1: Summary Statistics by Sample Group

	Matched	PKK	Random	Siblings
Population				
Birth City	102.83 (105.08)	103.18 (105.57)	148.15 (235.33)	94.63 (94.02)
Address City	1,811.50 (2,810.41)	2,125.52 (3,155.65)	2,393.35 (3,046.44)	2,006.81 (3,073.26)
Neighborhood	10.77 (17.83)	12.26 (19.96)	10.80 (17.93)	12.40 (20.90)
Demography				
Household Size	5.62 (4.00)	6.22 (3.75)	3.84 (2.70)	8.00 (4.05)
Gender	0.17 (0.38)	0.18 (0.38)	0.51 (0.50)	0.35 (0.48)
Age	29.01 (7.88)	28.82 (7.48)	43.42 (16.13)	29.36 (7.11)
Address Cell				
HDP Voteshare	51.44 (37.13)	54.75 (36.09)	15.09 (24.06)	61.83 (33.78)
NTL (2012)	31.30 (20.59)	32.45 (20.69)	34.84 (20.87)	32.11 (21.03)
NTL Change	12.42 (8.74)	12.53 (8.55)	12.53 (8.88)	12.49 (8.51)
Birth Distance	0.32 (0.47)	0.32 (0.47)	0.23 (0.35)	0.30 (0.46)
Count	44811	729	18942	1701

This table reports the mean and standard deviation (in parentheses) for the variables used in this analysis. The units for the population variables are thousands. "Neighbourhood" refers to the neighbourhood of residence. "Household Size" refers to the household of residence. "Address Cell" variables describe the characteristics of the cell of residence. "NTL" is shorthand for Nighttime Lights. Birth Distance indicates the euclidean distance in thousands of kilometers between an individual's location of birth and residence.

siblings of PKK recruits. The "Matched Sample" was constructed as follows: for each PKK recruit, 50 individuals were selected from the MERNIS database who were born in the same place, in the same year, and who were of the same gender as the recruit. The selection of individuals who are so demographically similar to the PKK members themselves enables a significant degree of latent variation to be accounted for; two women from the same village who were born in the same year are likely to face similar economic pressures, have similar ethnic identities, and perhaps even share political or familial ties.

1.5 Analysis

To explore the general characteristics associated with PKK membership, a simple logistic regression of the following form is estimated:

$$Y_i = \beta_0 + \beta_1 X_i + \mu_i + \epsilon_i \tag{1.2}$$

Where Y is a dummy variable denoting whether an individual i was a member of the PKK, X is the set of explanatory variables listed in Table 1.1, μ is a fixed effect for the city of residence, and ϵ is the error term. This model is estimated using logistic regression, and the results are shown in Table 1.2; The first column contains results derived from a comparison of deceased PKK recruits with the random sample from the MERNIS database, while the second column utilizes the matched sample.

Table 1.2 allows for a precise estimation of the demographic differences between PKK members and the general population. Compared to the random sample, PKK members are significantly more likely to be younger, male, and come from larger households: the likelihood of joining the PKK decreases by 9.6% per additional year of age, increases by 9.4% for every additional household member, and is 79.2% higher for men than it is for women. These results closely match established findings on the demographic characteristics of militants, but provide a higher degree of precision in the estimates of these relationships.

Furthermore, the fact that these results conform to expectations provides a useful form of validation when exploring themes that have less empirical evidence, such as the relationship between insurgency and migration. Militants tend to settle in areas that had higher vote shares for the main Kurdish party (the HDP) in the 2015 election. Because the matched sample was generated by selecting individuals with the same birth year, birthplace, and gender as the PKK sample, the coefficients for these variables are predictably insignificant in column 2. However, even in the matched sample, there is a strong positive association between household size, HDP vote share, and PKK membership.

The set of migration-related variables adds a further level of detail. City-of-residence fixed effects control for latent variation in recruitment rates related to the specific characteristics of a given city. The coefficients of these fixed effects in the random sample model are themselves informative, as they identify localities that have a significantly higher proportion of PKK recruits; these include many urban centers in the South and East of the country such as Tunceli, Hakkari, and Van. Despite

Table 1.2: Logistic Regression Estimates for Characteristics Associated with PKK Membership

	Random Sample	Matched Sample
Nighttime Lights	0.00	-0.00
	(0.01)	(0.00)
Nighttime Lights Change	0.00	0.01
	(0.01)	(0.01)
HDP Voteshare	0.01^{***}	0.01^{**}
	(0.00)	(0.00)
Household Size	0.09^{***}	0.05^{***}
	(0.01)	(0.01)
Neighbourhood Population	0.00	0.00
	(0.00)	(0.00)
Birthplace Population	0.00	-0.00
	(0.00)	(0.00)
Distance Traveled	0.81^{**}	0.88^{***}
	(0.29)	(0.26)
Gender	-1.57^{***}	0.07
	(0.11)	(0.10)
Age	-0.10^{***}	0.00
	(0.01)	(0.01)
Conflict Proximity	0.38^{***}	0.02
	(0.08)	(0.07)
Birth Latitude	-0.40^{***}	-0.04
	(0.06)	(0.06)
Birth Longitude	0.08^{*}	-0.10^{**}
	(0.03)	(0.03)
AIC	3640.05	7523.29
BIC	4381.42	8334.84
Log Likelihood	-1726.03	-3668.64
Deviance	3452.05	7337.29
Num. obs.	19671	45540

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

This table reports estimates from logistic regressions in which the dependent variable in all cases is a binary indicator of PKK membership. The "Random Sample" column compares PKK recruits to a random sample of Turkish citizens, while the "Matched Sample" column compares them to individuals with similar observed demographic characteristics (age, gender, origin). Both models include city-of-residence fixed effects, though coefficients are not reported.

the inclusion of the city-of-residence fixed effect and controlling for longitude and latitude– and therefore, that individuals born farther afield must travel farther– PKK members tend to travel greater distances than individuals in both the random and the matched sample. This result dovetails with the significant positive relationship between proximity to conflict and the likelihood of joining the PKK present in the random sample. Compared to the random sample, PKK members tend to be born closer to conflict events and move farther afield than the average Turkish citizen, even when controlling for cardinal position and ultimate destination.

Null results are also informative: there is no relationship between either the level or growth of nighttime lights and the likelihood of joining the PKK. In the "Youth Bulge" literature, frequent claims are made about overcrowded peri-urban areas becoming sites of conflict. Yet the results in Table 1.2 suggest that a PKK recruit is no more likely to live in an urban or peri-urban area than the average Turkish citizen.

The following five subsections explore the above results in greater detail. Section 1.5.1 compares country-level trends in the migration of militants and random citizens to urban areas. Section 1.5.2 examines city-level characteristics of the destinations chosen by militants. Section 1.5.3 observes spatial clustering at the neighbourhood-level and analyzes potential peer effects. Section 1.5.4 examines intra-household dynamics in the recruitment of children, including the role of household size and birth order. Section 1.5.5 introduces evidence suggesting the importance of political grievances in the process of insurgent recruitment.

1.5.1 Country-Level Trends in Migration

The majority of studies linking conflict and urbanization treat the latter as a black box. Many simply compare the growth rate of cities across countries and correlate this with conflict (Esty et. al. 1998, Buhaug and Urdal, 2013). Yet people migrate for myriad reasons, many of which are of direct relevance to the grievances of insurgents. Two of the results from the regression analysis in Table 1.2 are of particular relevance to understanding the possible causes of migration among PKK recruits. The first is that despite controlling for city-of-residence fixed effects and the fact that most recruits come from Southeastern Turkey, there is a significant positive relationship between the proximity of an individual's birthplace to conflict events and the likelihood of joining the PKK. The second important result is that PKK members tend to move farther away from their birthplaces than both random Turkish citizens, and even those who are nearly demographically identical to themselves. Taken together, these findings could suggest conflict-induced displacement as a possible source of grievance. This section analyzes migration patterns to further investigate this mechanism.

The top panel of Figure 1.3 displays migratory patterns among a random sample of anonymized records from the MERNIS database. Arrows connect the location of birth and last registered address of an individual and are colored according to change in longitude, with red denoting Westward migration and green denoting Eastward migration. The long red arrows show a general pattern of Westward migration from the rural areas in Turkey's East towards major cities such as Istanbul and Izmir. Though less numerous, a number of long Eastward journeys can be seen in green, typically originating from major cities. In addition to these longer journeys, the shorter yellow arrows indicate a substantial amount of local migration, with individuals often staying within the same district. According to a random sample of over 19,000 individuals, the average Turkish citizen travels 231 kilometers from their birthplace.

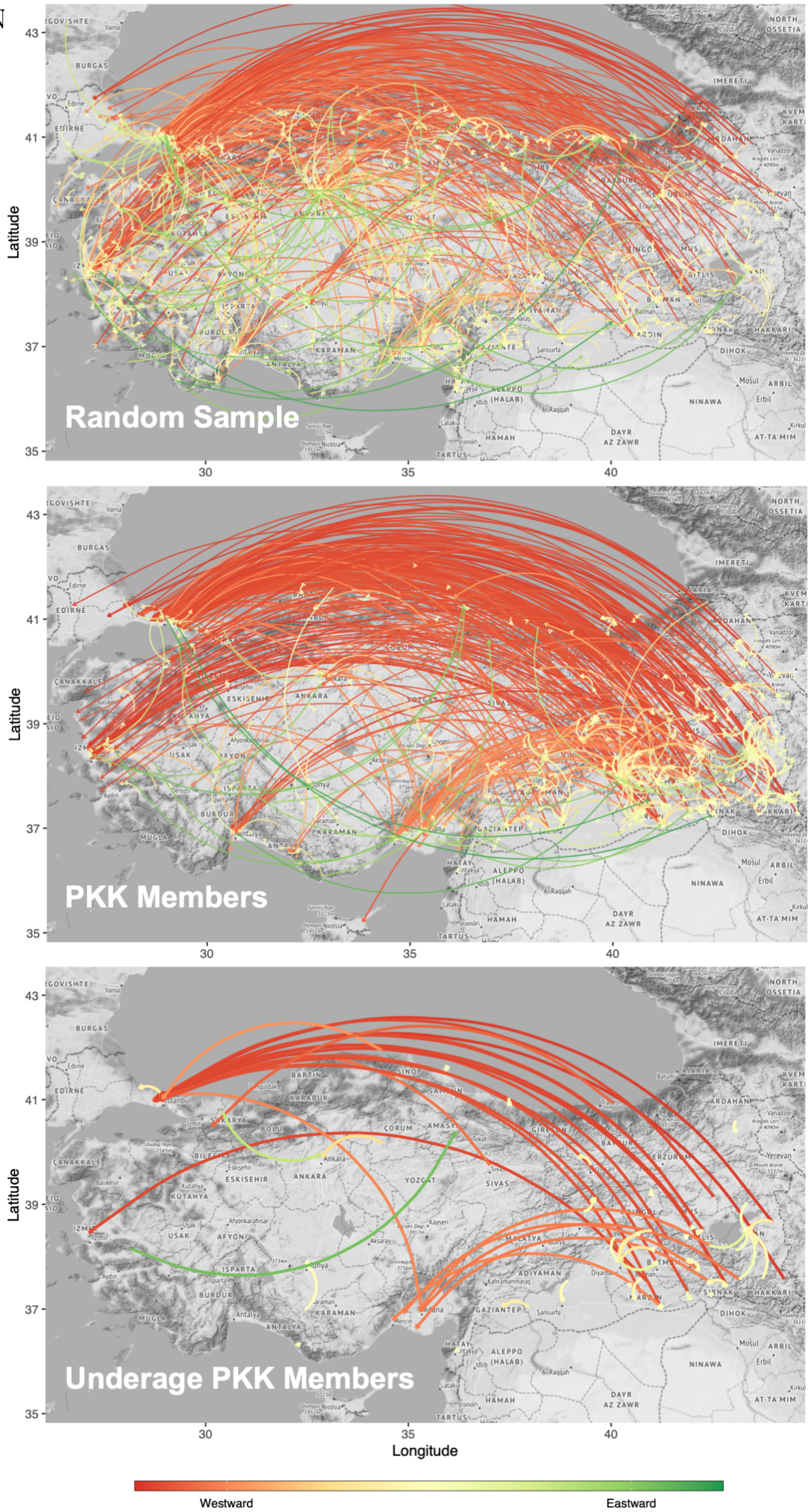
In contrast, the average PKK member travels 318 kilometers. The middle panel in Figure 1.3 displays the migratory patterns of PKK members. The PKK began as a "peasant movement" based largely in the rural agricultural villages of Turkey's Southeast. Yet even among those that did not leave the Southeast, the clusters of yellow arrows surrounding urban centers in the area suggest a substantial degree of local rural-urban migration. While the majority of recruits were born in the Southeast, most of them moved to major cities in the West prior to joining the insurgency. Though migration towards cities is a present in the random sample, there are substantial differences in the choice of destination. The random sample displays significant migration towards the capital city of Ankara, while the PKK sample shows virtually none. Conversely, many PKK members migrated from the rural Southeast towards Adana, a relatively unpopular destination in the random sample. In general, the choice of destination city among the PKK members is more concentrated than in the random sample. The three major urban destinations among PKK members are Istanbul, Izmir, and Adana, all of which have substantial Kurdish populations, which may reflect the tendency of individuals to migrate towards areas in which they may have familial or cultural ties. Section 1.5.2 explores differences in PKK members' choice of destination relative to the random and matched samples in greater depth.

Though the recruitment of child soldiers is often thought to be a mainly rural phenomenon (Boothby, Crawford, and Halperin, 2006), the proportion of urban recruits is higher for children than it is for the whole sample. Filtering the sample of PKK recruits to only include those who joined as children highlights this largely urban character of underage insurgent recruitment, shown in the bottom panel of Figure 1.3. Though the plurality of underage recruits were residents of Istanbul prior to joining the PKK, only one of them was born in the city itself. The vast majority were born in rural parts of the Southeast, and migrated to the city. In the Southeast, the recruitment of children is strongly characterized by local rural-urban migration from the villages surrounding the cities of Diyarbakir and Van.

While the random sample of Turkish citizens displays a strong general trend of Westward migration to urban centers, migratory patterns among PKK recruits are largely characterized by outmigration from the rural Southeast towards both local urban centers and major western cities. Though the causes of outmigration from the Southeast are manifold and include economic motivations, the depopulation of the rural Southeast is primarily attributed to the brutality of the Turkish government's counterinsurgency tactics in the 1990s, which included the burning of over 3,500 Kurdish villages, extrajudicial killings, torture, and forced disappearances (Van Etten et. al., 2008).

In an ethnographic study of Kurdish migrants residing in the Kanarya neighbourhood of Istanbul, Kilicaslan (2015: 163) found forced migration resulting from state violence to be the norm among local residents. All of the respondents were born in villages in the Southeast, and fled to Istanbul after either directly suffering violence at the hands of the state, facing the destruction of their homes or property, or fearing impending retaliation from paramilitary village guards. One respondent in the study narrated his reasons for leaving the Southeastern town of Cizre as follows: "In autumn at the end of 1993, guerrillas killed the father of the head of the village guard X. After this incident, the village guards got ill tempered and drew a line. They said "either these people will be with us or they will all migrate". The population of Cizre

Figure 1.3: Migratory Patterns of Random Turkish Citizens and Militants



These maps display arrows linking locations of birth and residence for random Turkish citizens, PKK members, and underage recruits. Arrows are colored by the cardinal direction of travel.

shrank from eighty thousand to ten thousand. Of course, some joined the village guards. In Cizre, those who did not take sides could not make a living. After that [the death of X's father] the number of village guards increased. After that, conflicts arose frequently. Our close friends became village guards, and they began to oppress us. Every night, they opened fire to frighten us."

This account seems to suggest that retribution for PKK attacks was exacted indiscriminately from the civilian population of the areas surrounding the attack. Beyond intimidation and the deliberate economic marginalization of those who refused to join the state-backed village guard paramilitaries, other respondents narrated far more direct forms of coercion:

"The day after the armed conflict, the village was in chaos. They burnt the houses and the animals, and they also killed people. Seven of my relatives were killed. We saw with our own eyes that they burned the houses by throwing something white on them".

Though several respondents described attempting to relocate to neighbouring villages, the pervasiveness of conflict across the region ultimately informed their decision to move to Istanbul. This qualitative evidence suggests that those who migrate due to their proximity to conflict are forced to travel longer distances, and carry with them a powerful source of grievance.

Though neither the results from Table 1.2 nor the analysis of migratory patterns can provide conclusive evidence that PKK recruits are disproportionately conflict-induced migrants, they do suggest that PKK members are born significantly closer to conflict than the average Turkish citizen, and migrate longer distances than their demographic peers. These patterns align closely with ethnographic literature on the the characteristics of conflict-induced displacement in Turkey. While the majority of the literature on conflict and urbanization focuses on the growth of the latter, this analysis suggests that the forces driving urbanization may themselves also drive insurgent recruitment.

1.5.2 City-Level Trends in Urban Recruitment

Having examined the relationship between migratory push factors and insurgent recruitment, this sub-section examines the characteristics of migrants' destination cities and their relationship with insurgent recruitment.

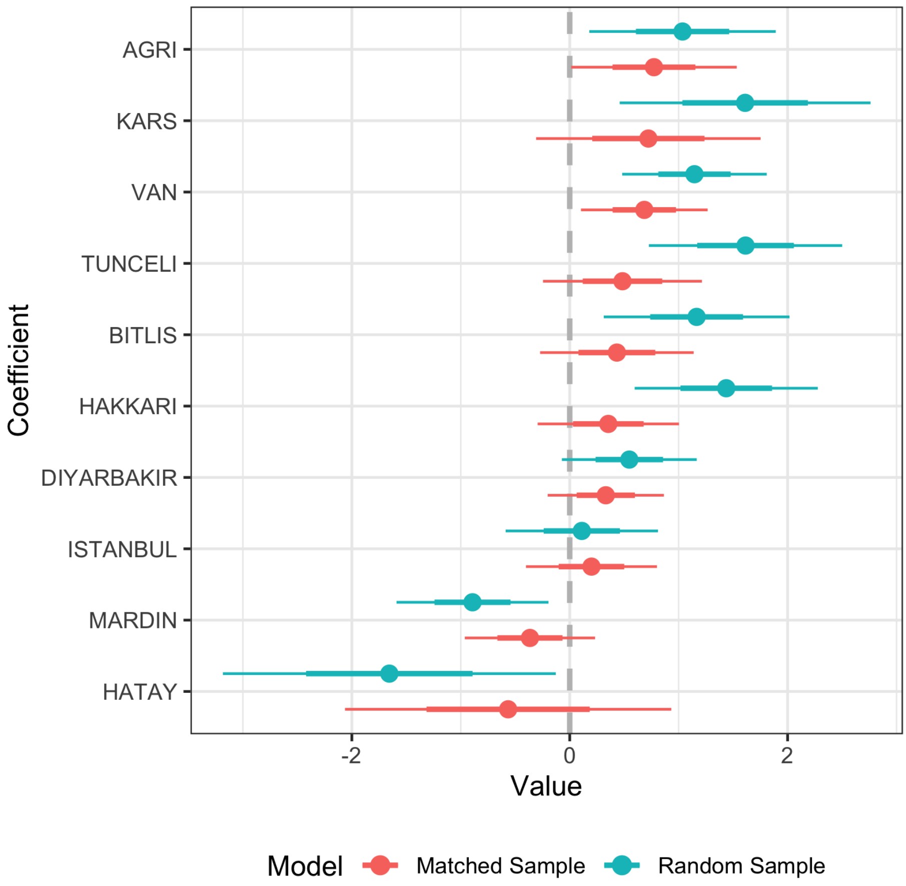
Figure 1.4 plots selected coefficients of the city-level fixed effects in Table 1.2, highlighting cities that have a disproportionately high or low number of PKK recruits. Several Turkish cities are significantly more likely to act as destinations for PKK recruits when compared to the general population. Intuitively, the effect size is diminished when the matched sample is used; people from the same villages and with the same demographic characteristics as PKK recruits make more similar decisions in terms of their city of destination than do random Turkish citizens. There is one exception: the city of Van in Southeastern Turkey, which remains a significantly more frequent destination for militants even when the matched sample is used.

Having identified a city in which insurgent recruitment is heightened, mechanisms linking urbanization and conflict in the literature can be tested at an unprecedentedly detailed scale. Gizelis, Pickering et. al. (2021: 1) "[...] expect population pressure to have the most profound effects on social unrest in peri-urban areas, meaning the urban outskirts." Rather than being motivated by political grievances, this literature frames conflict as an outgrowth of poverty and competition over scarce resources on the urban fringes.

To assess whether recruitment in Van is concentrated in peri-urban areas, Figure 1.5 displays the relationship between nighttime lights growth and the spatial distribution of PKK recruits' addresses.

Red indicates growth in nighttime lights between 1992 and 2012. Blue indicates little change in nighttime lights, which includes both rural areas which were dark and remain dark, as well as the city centre– which was bright and remains bright. The black circles represent militants who were born and raised in Van, while the crosses represent militants who migrated to the city. There are no obvious differences in the spatial distribution of local and migrant militants in Van. However, there are four instances in which a cross and a circle are directly superimposed, meaning a militant migrated to the same address as another militant who was born and raised in the city.

Figure 1.4: City Fixed Effects on Likelihood of PKK Membership



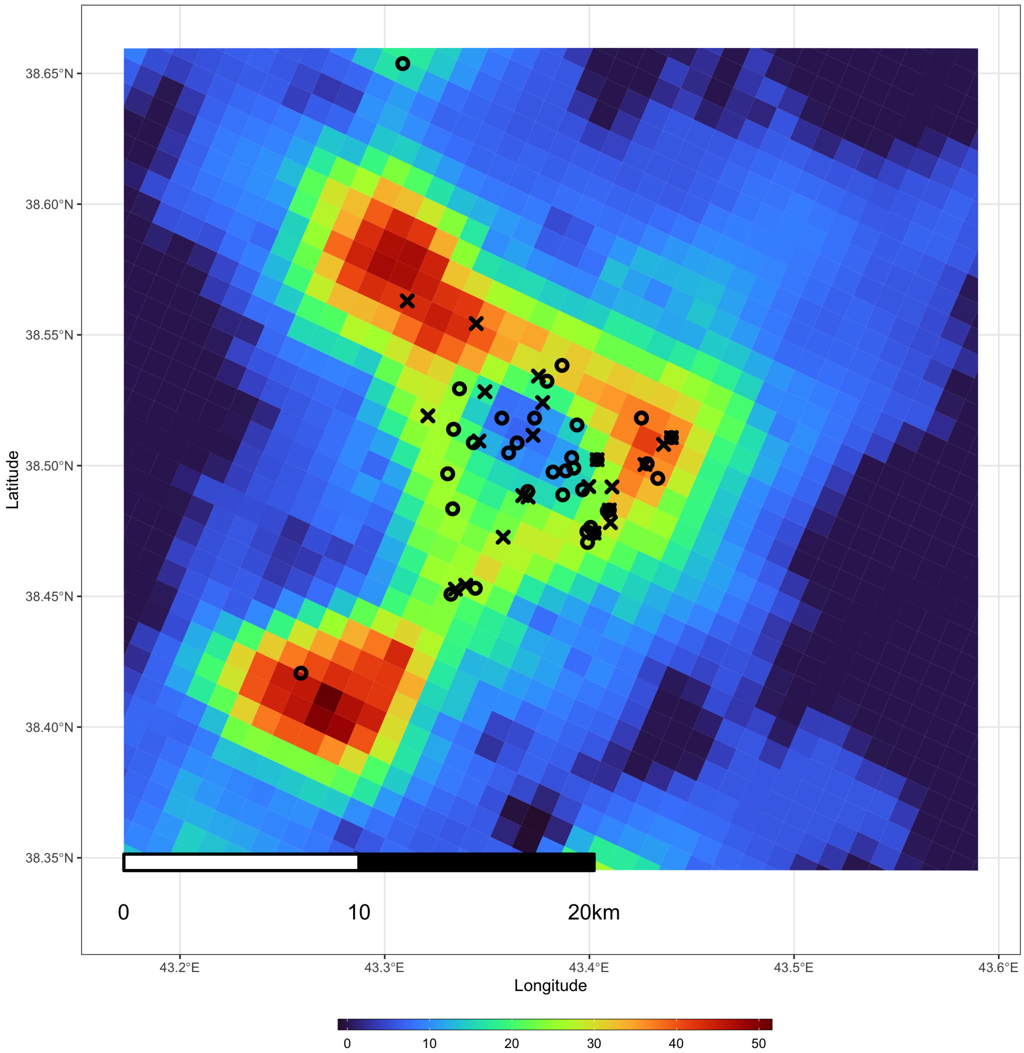
This plot shows the coefficients for city-level fixed effects from Table 1.2.

This seems highly suggestive of social or familial effects, which are explored in greater depth in Section 5.4. In Van, there does not appear to be a greater concentration of recruits in the peri-urban ring surrounding the city.

Though the regression results do not indicate that Istanbul is a significantly more popular destination for PKK recruits than for the general population, it remains the country's largest city and is home to more insurgents than any other city. A closer look at urbanization and insurgent recruitment in Istanbul reveals similar trends to Van, shown in Figure 1.6.

Clusters of insurgents can indeed be seen in certain peri-urban areas; the neighbourhood of Esenyurt in the West of the city experienced moderate nighttime lights growth (an increase of around 20 radiance units), and was home to a cluster of nine insurgents. Much like the aforementioned Kanarya neighbourhood– which happens to be only a

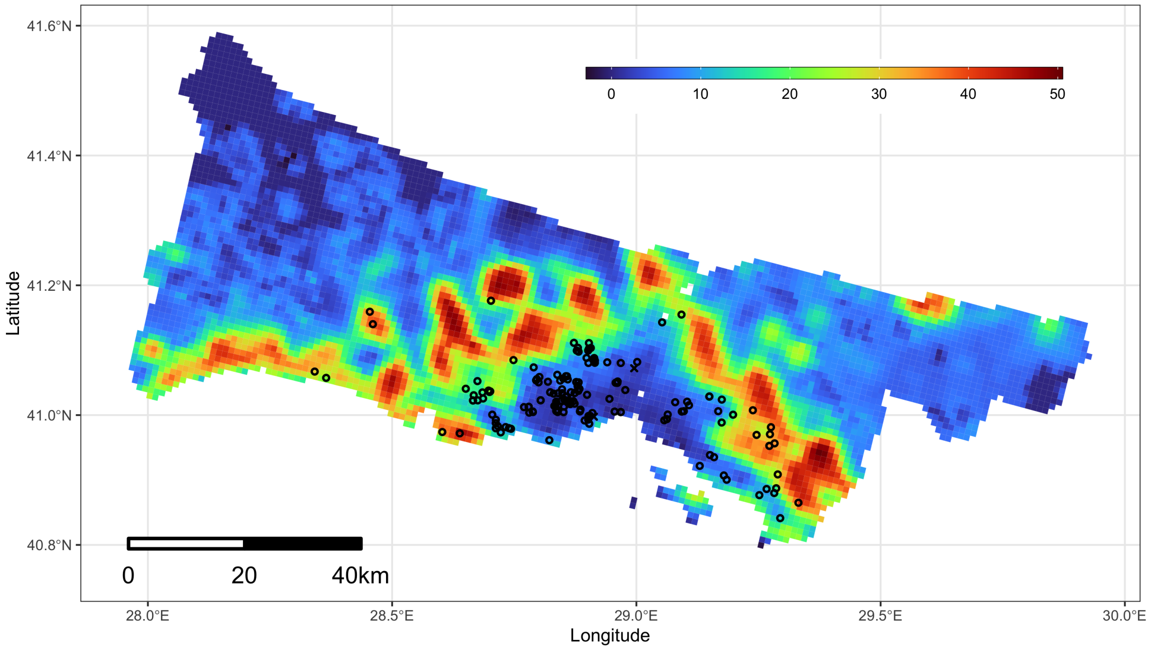
Figure 1.5: Urban Growth and Recruitment in Van, 1992-2012



The underlying raster shows the growth in Nighttime lights between 1992-2012 in the city of Van at the level of 1km^2 grid cells, with red denoting new urban growth. Black crosses represent militants who migrated to the city, while circles represent those who were born there.

few kilometers away– Esenyurt became a popular destination for forcibly displaced Kurds, and many of its residents consequently harbor strong grievances against the Turkish government (Ahmetbeyzade, 2007). Yet the majority of recruits do not live in these new peri-urban areas, contrary to what is hypothesized in the literature on conflict and urbanization. In Istanbul, insurgents largely live in neighbourhoods that were already well established (i.e. luminous) by the time they arrived. Strikingly, out of over 100 recruits who lived in Istanbul prior to joining the PKK, all but two migrated to the city from elsewhere.

Figure 1.6: Urban Growth and Recruitment in Istanbul, 1992-2012



The underlying raster shows the growth in Nighttime lights between 1992-2012 in the city of Istanbul at the level of 1km^2 grid cells, with red denoting new urban growth. Black crosses represent militants who migrated to the city, while circles represent those who were born there.

Overall, both figures suggest that relatively few recruits were living in the peri-urban areas: most lived in close proximity to each other in older neighbourhoods in the city's center. Indeed, the regression results in Table 1.2 show no relationship between either the level or growth of nighttime lights and the likelihood of being an insurgent.

The lack of association between nightlights growth and insurgent recruitment does not itself preclude the possibility that urban slums play a role in the recruitment process. Urban growth in Turkey has been characterized by a type of informal settlement known as the *gecekondu* ("night building"), involving dwellings that are erected overnight. Rather than forming a ring around the outskirts of cities, many of these settlements are spread throughout more central urban areas. Because these informal settlements do not contain registered streets and the dwellings themselves have no precise address, they can be readily identified in the MERNIS database. Individuals who live in these informal settlements have the phrase "Küme Evleri" ("cluster of homes") instead of a street name. In the random sample of Turkish citizens, 3.3% of individuals were identified as living in informal settlements. Contrary to the supposition in the literature that urban slums constitute breeding grounds for violent extremism (Patel and Burke, 2011), the proportion of PKK members living in informal settlements is just 0.2% higher than that of the general population, at 3.5%.

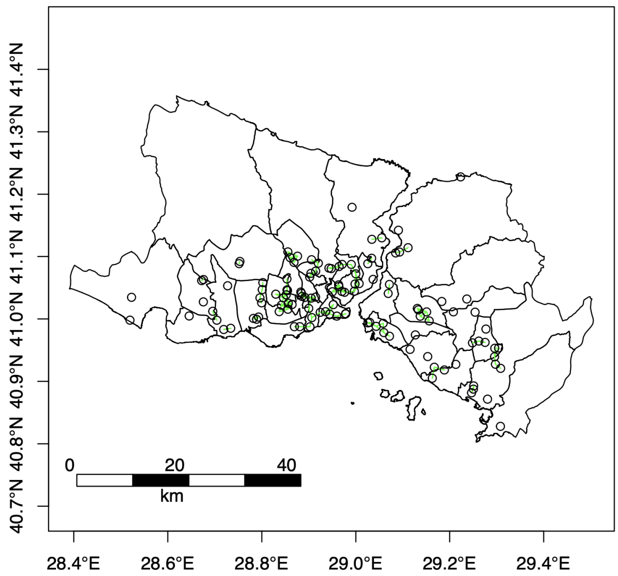
This sub-section finds that insurgents are no more likely to live in urban slums than the average Turkish citizen, and have tended to migrate in tight spatial clusters to established neighbourhoods in the inner city rather than to growing peri-urban areas. Indeed, there were several instances in which militants migrated to the same address as another militant born and raised in the city. Both of these trends could suggest social and familial factors as being involved in the process of insurgent recruitment. The following sub-section interrogates the possibility of peer effects in the process of urban insurgent recruitment through a neighbourhood-level analysis of spatial clustering in Turkey's largest city.

1.5.3 Neighbourhood-Level Peer Effects and Spatial Clustering in Istanbul

Though there is ample qualitative evidence to suggest that social ties are a crucial factor in insurgent recruitment (Brett and Specht, 2004), there is very little quantitative research into the nature of these dynamics. Helmers and Patnam (2014: 67) use "spatial peer interaction, defined as a child's nearest geographical neighbours" to assess the impact of peer effects on learning outcomes. Though well established in economics, no such research design has been possible in the study of militancy. Through a comparative analysis of spatial clustering, this sub-section assesses the role of geographic– and therefore, social– proximity in the process of insurgent recruitment.

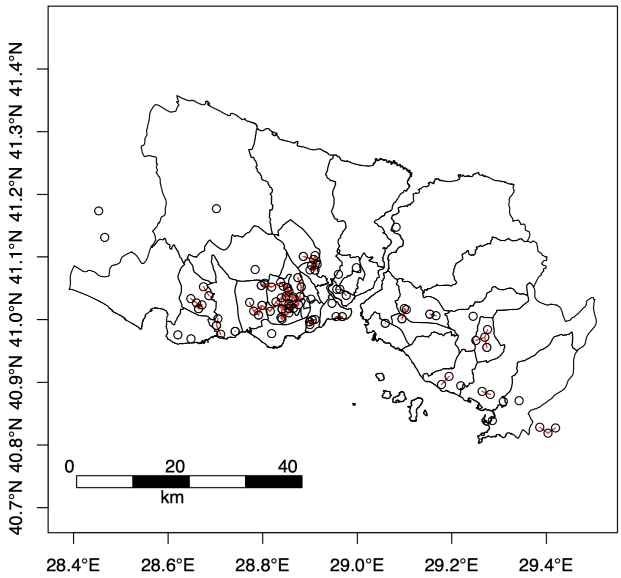
129 of the PKK recruits that were matched with records in the MERNIS database had addresses in Istanbul. Figure 1.7 shows the spatial distribution of 129 random addresses across Istanbul taken from the MERNIS database. Addresses are linked by a green line if they are within a 2 kilometer radius of each other. Though the majority of addresses are located West of the Bosphorus, the addresses are relatively spread out across the city. Many of the addresses have no neighbours within 2km, and those that do are arranged into smaller clusters of 2-5 addresses.

Figure 1.7: Random Sample 2km Neighbours



This plot shows a random sample of addresses (circles) in Istanbul, which are connected with a green line if they are within 2km of each other.

Figure 1.8: PKK Members 2km Neighbours



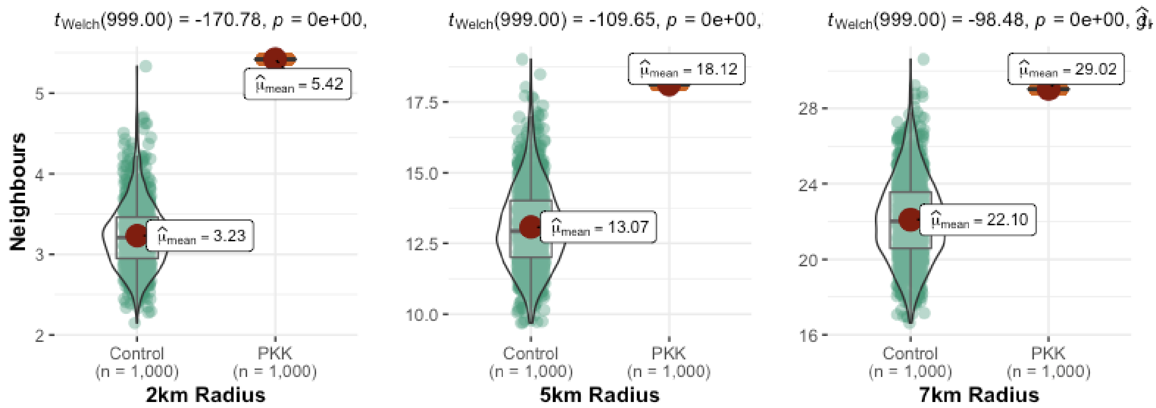
This plot shows significant spatial clustering among PKK members in Istanbul compared to the distribution of random addresses on the left.

When the same procedure is applied to a sample of PKK members' addresses (Figure A1), dense geographic clusters arise. Recruits appear to be heavily concentrated on the European side, particularly in the neighborhoods of Bahçelievler and Bağcılar. Compared to the random sample, very few of the addresses in this sample have no neighbours within 2km. Rather than exhibiting a large number of small clusters, this sample exhibits a small number of very large clusters. Figure **??** restricts the sample of recruits to those who joined prior to their 18th birthday, displaying an even stronger geographic concentration: all of the underage recruits in the sample come from a handful of neighbourhoods West of the Bosporus.

These plots suggest that PKK members tend to live significantly closer to each other than do random people. Figure 1.9 computes the average number of neighbours at various distance thresholds by drawing 1000 random samples of 129 addresses in Istanbul from the MERNIS database. The resulting mean is compared to the number of neighbours among the 129 PKK recruits from the city.

The leftmost distribution indicates that when 129 addresses are chosen at random, the average Istanbulite will live within 2 kilometers of 3.23 of these addresses. In contrast, the average PKK member lives within 2 kilometers of 5.42 other PKK

Figure 1.9: Number of Neighbours Within a Given Distance Radius in Istanbul

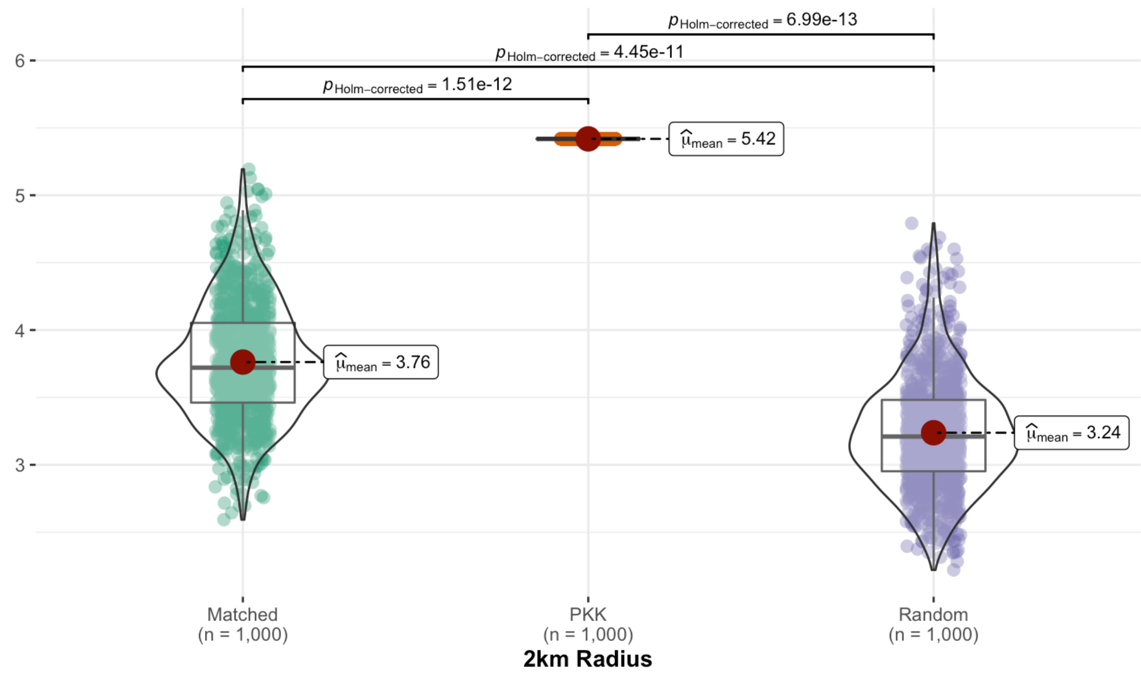


Each green point represents the number of individuals out of a random sample of 129 Turkish citizens who live within a specified distance radius of each other (2km, 5km, and 7km). Each subplot compares the number of neighbours within a given distance radius between 1,000 random samples of Turkish citizens and a sample of PKK members, for the city of Istanbul.

members. Of the 1000 random samples drawn in the leftmost plot, not one displayed geographic clustering as extreme as the PKK sample. This discrepancy attenuates slightly as the distance radius is expanded; average Istanbulites will have 13.07 neighbours within 5km, compared with 18.12 among the PKK members. At 7km, the number of neighbors increases to 22.10 and 29.02, respectively. Though there appears to be spatial clustering of PKK members compared to the random sample of Turkish citizens, this could simply be due to "cumulative causation": the process by which individuals migrate to areas in which there are existing familial or diasporic connections (Massey, 1990). Because the vast majority of PKK members are ethnic Kurds, the fact that PKK members appear to live close together could simply be due to the fact that Kurds will move to neighbourhoods in which there is an established Kurdish community. Indeed, a leaked U.S. State Department cable found that "a typical Kurdish rural-to-urban migrant is not, and cannot be, 'an individual actor.'", and that "maintaining one's 'Kurdishness' is a practical necessity for access to housing and employment and social acceptance." (Wikileaks, 2003).

The matched sample can be used to control for ethnic differences: for each PKK member in the sample, an individual is chosen with the same birth year, gender, and birth city. Beyond the fact that people born in the same village are likely to be of the same ethnicity, selecting individuals of the same age and gender as insurgents

Figure 1.10: Matched Sample Nearest Neighbours in Istanbul



Each purple point represents the number of individuals out of a random sample of 129 Turkish citizens resident in Istanbul who live within a 2 kilometer radius of each other. Each green point represents the same procedure, but applied to the sample of individuals who were demographically matched with PKK militants. 1000 samples are drawn for the random and matched groups, and compared to the mean number of neighbours of neighbours among the 129 PKK members resident in Istabul, shown as an orange bar.

further controls for potentially confounding factors that might affect an individual's location of residence such as gendered differences in intergenerational cohabitation, house leaving age, and labour market participation. Figure 1.10 compares the number of neighbours within a 2km radius between PKK members, the random sample of Turkish citizens, and the matched sample.

The matched sample displays a slightly higher level of spatial clustering than the random sample, potentially reflecting the tendency of individuals to migrate to areas in which they have diasporic connections. On average, individuals in the matched sample have 3.76 neighbours within 2 km, compared to 3.24 for random Turkish citizens. However, even when compared with individuals of the same age, gender, and provenance, PKK members still live in significantly closer proximity to each other; as such, this spatial clustering cannot simply be attributed to demography or ethnic identity.

Though impossible to prove conclusively, local social networks could explain the dense spatial clustering of PKK recruits in urban areas. In a study of the PKK, Tezcür (2016: 249) finds that "Individuals embedded in dense social networks with linkages to an insurgency are likely to join that insurgency". Five of the PKK recruits in the current sample lived in the Kanarya neighbourhood of Istanbul, which according to the ethnography cited in Section 5.1 is home to large numbers of Kurds who were forcibly displaced by conflict (Kilicaslan, 2015). Two of the recruits lived on the same street, were born one year apart, and despite having each traveled nearly 1,200km from their home villages in Southeastern Turkey, were born less than 20km from each other. Beyond direct peer effects, social and physical proximity to other PKK members could logistically facilitate an individual's decision to join the insurgency by providing knowledge of the recruitment process.

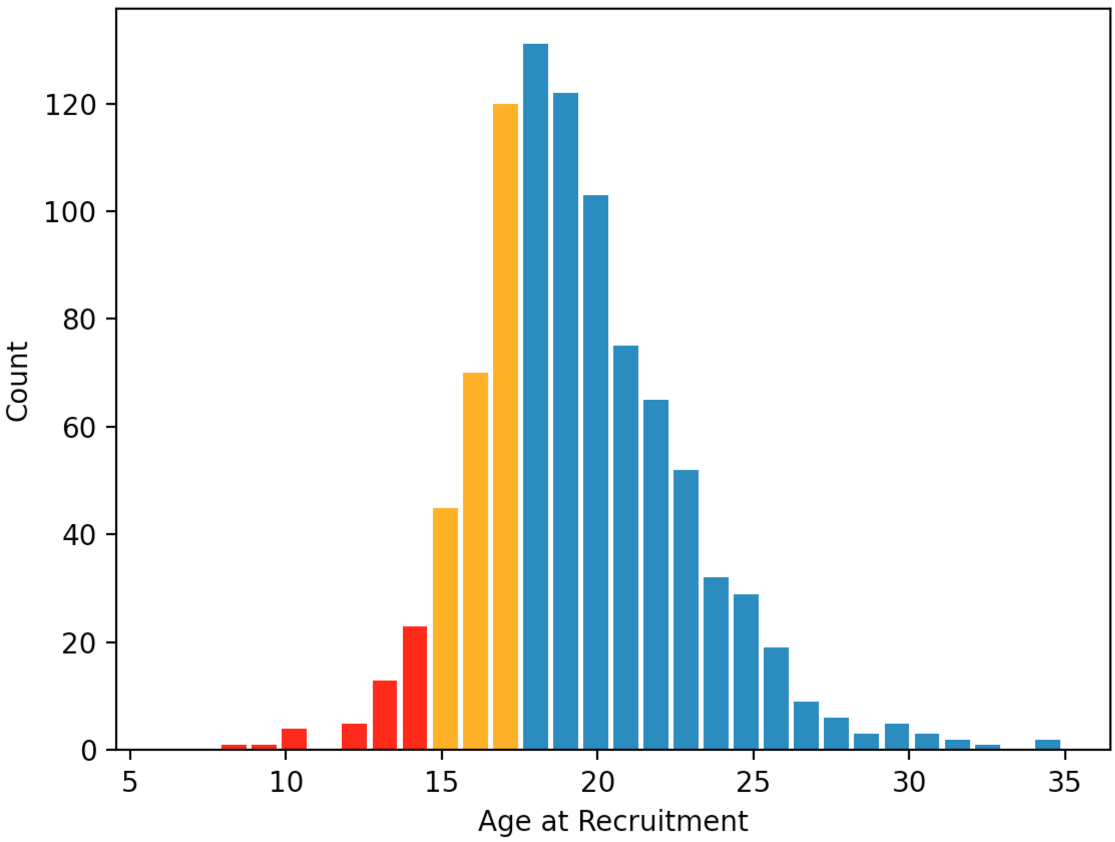
The literature on urbanization and conflict rarely addresses the social dimensions of militancy, focusing instead on factors such as poverty, overcrowding, and employment opportunities in urban and peri-urban areas (Esty et. al., 1998; Buhaug and Urdal, 2013; Gizelis, Pickering et. al., 2021). Yet considering that 127 out of the 129 recruits in Istanbul migrated to the city, the spatial clustering of militants suggests that these social dynamics are an intrinsic part of both the process of urbanization and its relationship with militancy.

1.5.4 Household-Level Dynamics in the Recruitment of Children

The previous sub-section identified the likely presence of neighbourhood-level peer effects in the process of insurgent recruitment. Indeed, in the entire sample of PKK recruits, 18.8% lived on the same street as at least one other militant. Of these individuals, nearly half (7.5% of the full sample) were from the same household. Household-level factors such as kinship ties (White and Guest, 2003), household size (Blake, 1989), and birth order (Doyle et. al. 2020) are recognized as being important mediators of key economic outcomes for young people. Though qualitative studies affirm the importance of these factors in the process of insurgent recruitment (Brett and Specht, 2004; Tezcür, 2016), quantitative empirical research thereon is lacking. This sub-section fills this gap. Because the effect of familial factors is likely to disproportionately affect young people, I first provide an overview of the age and demographic structure of the sample of PKK recruits. Following this, I explore regression results in greater detail and conduct new analysis on intra-household dynamics including household size and birth order.

The demographic variables in Table 1.2 suggest that PKK recruits are significantly younger and more likely to be male than the average Turkish citizen. Though a greater amount of information was available for individuals whose obituaries could be linked with the MERNIS database, the obituaries themselves contain enough information (year of birth and year of recruitment) to plot the distribution of ages at which individuals join the PKK, shown in Figure 1.11.

Figure 1.11: Age-at-Recruitment Distribution



This distribution shows the age at recruitment derived from PKK obituaries. Recruits under the age of 18 are shown in orange, and those under 15 are shown in red. The mean age at which individuals join the PKK is 19.

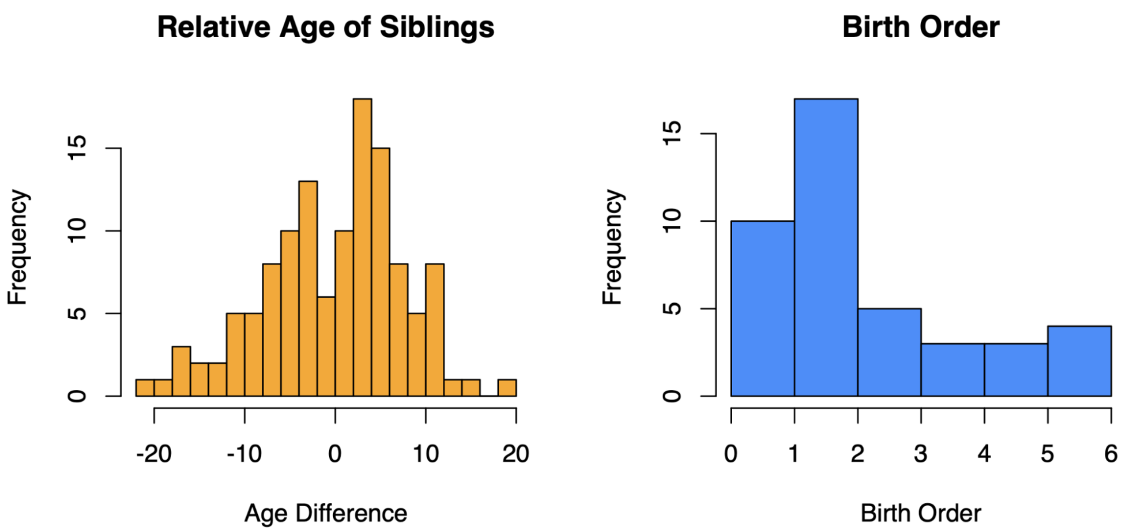
Recruits under the age of 18 are shown in orange, and those under the age of 15 are shown in red. The plot above shows that the average age at which individuals join the PKK is 19, and that a significant proportion of their cadres were under the age of 18. By the PKK's own admission, there were 278 child soldiers killed between 2001 and 2012, constituting roughly 30% of the sample for which there is complete information. This is likely to be an underestimation for two reasons. First, because recruits who joined as children but are still alive do not have obituaries. Second, because it relies partially on information reported by the PKK, who have an incentive to minimize the number of child soldiers they publicly admit to having recruited. Furthermore, because the PKK stopped reporting birth and recruitment dates in 2012, this method can't be applied to the 2,230 obituaries listed since then.

Given the significant proportion of PKK recruits that are under the age of 18, the absence of research into the relationship between insurgent recruitment and factors relevant to adolescent development constitutes an important gap in the literature. There is ample evidence from Sociology and Economics establishing an inverse relationship between the number of siblings in a family and outcomes such as labour market performance and education (Breland, 1979; Downey, 1995; Blake, 2020). The proposed mechanism for this relationship is often termed "resource dilution", whereby a child's achievement is linked to parental investment which is divided across siblings (Downey, 2001).

The regression results reported in Table 1.2 suggests that such a mechanism may also affect insurgent recruitment: there is a strong positive correlation between household size and PKK membership for both the random and matched samples. The median household size is 3 for the random sample, 5 for the matched sample, and 6 for PKK recruits. Again, the use of the matched sample precludes the relationship between household size and insurgent recruitment being attributable to the fact that Kurdish families tend to be larger; even relative to those with the same demographic characteristics, PKK recruits come from significantly larger families. To explore this finding in greater depth, I conduct an intra-household analysis of birth order and the recruitment of children.

For each underage PKK member, siblings were identified as individuals sharing the same address, last name, and both parents' names as the recruit. Upon identification, only these individuals' dates of birth were exported from the database and assigned a numerical family identifier linking them to the corresponding recruit. In Figure 1.12, the histogram on the left shows the distribution of siblings' ages relative to the child in the same family that joined the PKK. On the right, the distribution of birth orders of child PKK recruits is shown. Both of these plots exclude recruits without siblings.

Figure 1.12: Relative Age and Birth Order of PKK Recruits with Siblings



PKK members who joined as children tend to be second in the birth order, typically 5 years younger than their siblings.

Though the relatively small sample size limits the feasibility of formal tests, the plurality of children who join the PKK are second-born children, by a substantial margin. The plot on the left suggests that many of those who join are 4-6 years younger than their siblings.

These trends align closely with the growing literature in economics on the effect of birth order on educational achievement, crime, and violent behaviour (Rahav, 1980; Black, Devereux, and Salvanes, 2005). The mechanism underlying the finding by Doyle et. al. (2020) that second-born children are significantly more likely to engage in delinquency is hypothesized to be greater parental investment in the first-born child. An ethnography of low-income Kurdish migrant families in Istanbul suggests an analogous process may govern the role of first-born children in the family. Bahar (2016) found that birth order was one of the main factors influencing parents' decision making related to questions of child labour: "As older children were the first ones to attain the minimum age mothers thought children could work, they were the first to be considered for work" (2016: 22). Among these families and in wider Kurdish culture, the first-born child carries significant responsibilities towards their siblings in terms of both care and income. The decision to join the PKK and leave one's family would entail an abdication of this responsibility for the first-born child, but not the second-born.

Though the literature on urban youth bulges and conflict emphasizes the role of employment and education in generating social unrest, there is a general lack of attention to the intermediary factors that mediate this relationship. This sub-section provides empirical support for the notion that well-established mechanisms linking youth achievement with household-level variables such as family size and birth order apply to the process of insurgent recruitment.

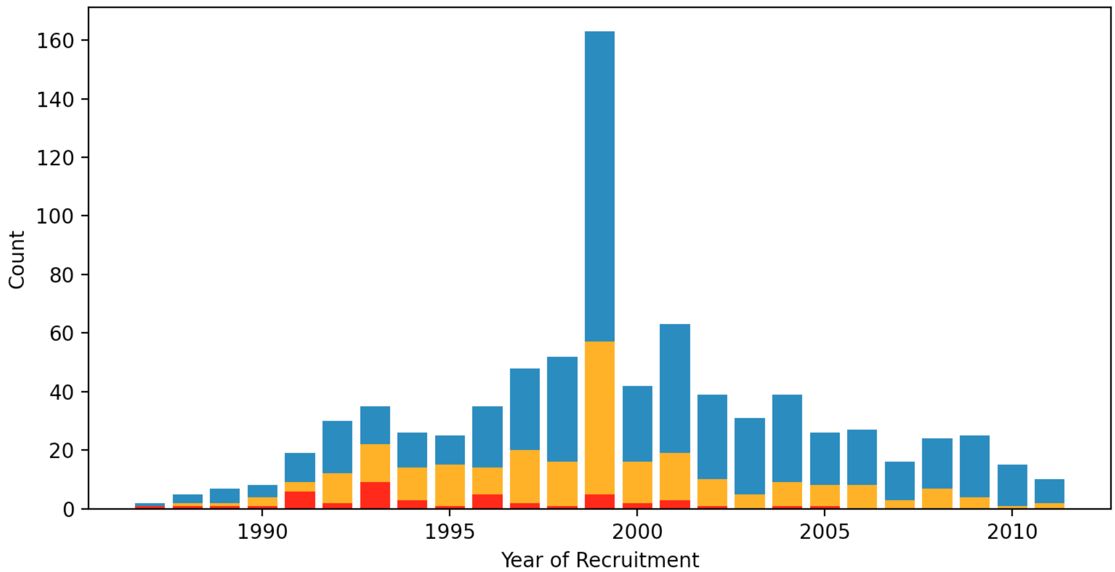
1.5.5 Individual-Level Political Grievances

Though social and economic factors do appear to play an important role in motivating young urban militants, the role of political grievances is explicitly downplayed in many of the most highly cited papers on insurgent recruitment (Collier and Hoeffler, 2004; Fearon and Laitin, 2003). Even Kurdish intellectuals and U.S. Diplomats view Kurdish separatism as being fueled by "Kurdish youth who have no particular political goal" (Wikileaks, 2010). This final sub-section interrogates these claims in reference to regression results and empirical trends in PKK recruitment that align with salient developments in Kurdish politics.

The results from Table 1.2 show a strong and consistent positive association between the vote share of the main Kurdish party (the HDP) and the likelihood of joining the PKK in both the random and matched sample. In the random sample, a 1% increase in HDP vote share in the square kilometer in which an individual resides was associated with a 1.01% increase in the odds of being an insurgent. Though the effect size is diminished to 0.61% in the matched sample, the result remains significant and therefore suggests that HDP voteshare is not simply proxying for areas with a large Kurdish population, instead signifying political engagement with the Kurdish movement.

Trends in recruitment dates provide further support the notion that political grievances are an important motivating factor. Figure 1.13 shows the number of recruits who joined the PKK in a given year. To visually assess the differential impact of political motivations across age groups, recruits are separated into three age groups: blue denotes adult recruits, orange denotes recruits under the age of 18, and red shows recruits under the age of 16.

Figure 1.13: PKK Recruitment by Year and Age Cohort



This figure shows PKK recruitment over time, with adult recruits shown in blue, those under the age of 18 shown in orange, and those under the age of 15 shown in red. A large spike in recruitment is evident in 1999, the year in which PKK founder Abdullah Öcalan was arrested and imprisoned.

The most salient trend in Figure 1.13 is the spike in recruitment in 1999. On February 15th of that year, the PKK's founder and leader Abdullah Öcalan was captured and sentenced to death. Kurds in over 20 countries staged protests and in several cases stormed embassies (Lyon and Uçarer, 2010). Several protesters were killed, and a Kurdish prisoner in Turkey self-immolated. Protests have been held in Turkey on February 15th every year since then. These heightened grievances are likely the cause of the spike in PKK recruitment observed in 1999, which saw more than three times as many recruits as the previous year.

An important caveat is that these are PKK-reported recruitment dates, and may thus be subject to bias. The PKK may have an incentive to over-report the number of individuals joining in this year to present an image of the organization as being highly politically motivated, or to convey that actions against its leaders will be met with surges in popular resistance. Whether this is the case remains unknowable from the data, but there is direct evidence to suggest that the arrest of Öcalan was an important motivating factor for some PKK recruits; a former member of the PKK stated in an interview "The arresting and imprisonment of Abdullah Öcalan was effective [sic] in my decision to join the PKK. Before going to the mountains, I first worked in the local institutions as a militia, after the arrest of Öcalan, I left the local institutions of the PKK, and joined the struggle in the mountains." (Aytekin, 2019: 72).

Disaggregating recruitment by age does seem to suggest heterogeneity in the relevance of political motivations, however. Though the absolute number of underage recruits increased in 1999, they made up a smaller proportion of the total number of recruits than in previous years, particularly children under the age of 16. Yet overall, insurgent recruitment is tied closely to both electoral outcomes and salient political developments, suggesting that the dismissal of political motivations found in much of the literature on militancy is unwarranted in the Turkish context.

1.6 Conclusion

The notion that large urban youth cohorts are one of the main drivers of civil unrest is widely held among policymakers and academics alike. Indeed, documents published Wikileaks indicate that the Kurdish insurgency was understood as resulting from a "youth bulge" by both Kurdish leaders and diplomats (Wikileaks, 2010). Yet in a review of the literature on the relationship between urban youth bulges and conflict, Cramer (2011: 2) concludes that "we still know too little empirically" to make any robust conclusions thereon. This problem is rooted in the inaccessibility of data required to test competing explanations, which requires individual-level observations of both insurgents and non-participants; such data is exceedingly rare, with one of the only exceptions being a study by Humphreys and Weinstein (2008) which involved large scale fieldwork in a conflict zone.

This paper creates a novel and unprecedentedly detailed dataset which allows for individual-level comparisons between militants, and both random and demographically matched samples drawn from the entire adult population of Turkey. I leverage this dataset to generate empirical insights that answer latent questions in the literature on youth, urbanization, and conflict.

While the majority of the literature on conflict and urbanization focuses on the quantity of the latter (Esty et. al. 1998, Buhaug and Urdal, 2013; Goldstone, 2002), the present analysis suggests that the forces driving urbanization can themselves also drive insurgent recruitment. A comparison of the migratory patterns of individual insurgents with a random sample of the Turkish population finds that PKK members are born significantly closer to conflict events than the average Turkish citizen, and migrate longer distances than even their demographic peers. These patterns align closely with ethnographic evidence on the characteristics of conflict-induced displacement.

Analysis of the city-level characteristics of the destinations chosen by migrating militants provides little empirical support for claims made in the literature that inner-city slums (Patel and Burke, 2011) and peri-urban areas (Gizelis, Pickering, et. al. 2021) are particularly conflict-prone. Rather than moving to the urban periphery or inner-city slums, militants migrate to established neighbourhoods in city centres, sometimes even migrating to the same address as another militant who was born and raised in the city.

These findings seem suggestive of peer effects. Despite ample econometric evidence linking peer interactions with key outcomes for young people (Burke and Sass, 2011; Zimmerman, 2003; Sacerdote, 2001; Trogdon et. al. 2008; Helmers and Patnam, 2014), quantitative research thereon in the realm of insurgent recruitment is absent. Though spatial proximity does not necessarily imply social proximity, an analysis of spatial clustering among militants at the neighbourhood-level suggests that peer effects may be present in the process of insurgent recruitment: militants are nearly twice as likely to live near other militants as are random Turkish citizens, and they sometimes even migrate together.

Indeed, a substantial number of PKK recruits come from the same household. Sociologists and economists have long observed relationships between familial factors such as household size, birth order and youth achievement (Breland, 1979; Downey, 1995; Blake, 2020; Doyle et. al. 2020). This study finds empirical support for the proposition that similar mechanisms may apply to the process of insurgent recruitment.

Finally, the contention that individual-level political grievances are irrelevant to the study of militancy (Collier and Hoeffler, 2004; Fearon and Laitin, 2003) is called into question via the identification of a significant relationship between the vote share for the main Kurdish party and insurgent recruitment. This is bolstered by the observation of a threefold spike in recruitment coinciding with a salient development in Kurdish politics: the capture of PKK founder Abdullah Öcalan.

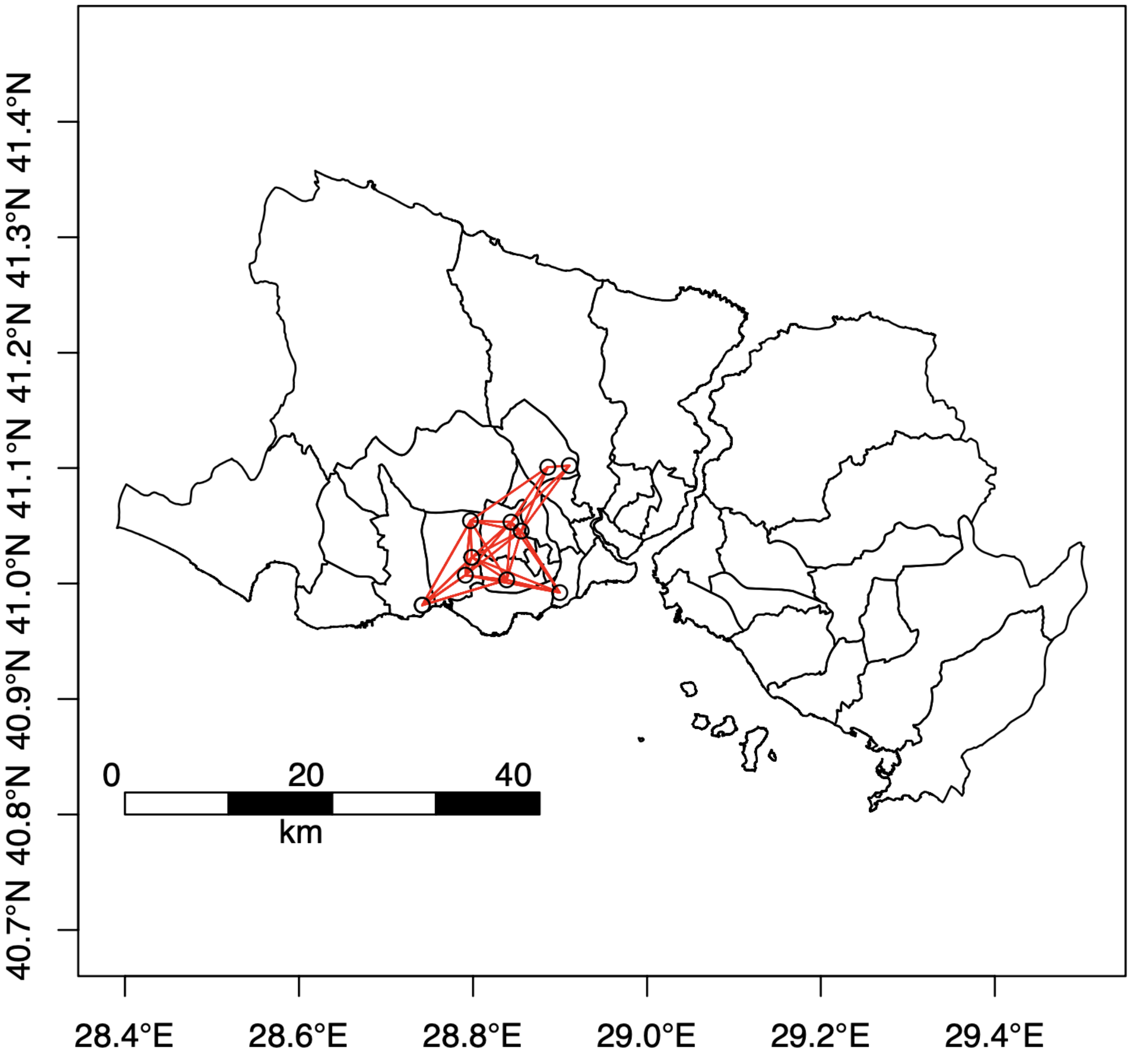
There are important limitations to this analysis, however. The first is that the dataset of PKK members is likely biased towards lower-ranking individuals, and may thus not represent the experiences of higher-ranking individuals. Secondly, despite the high spatial resolution of the MERNIS data, there is no temporal dimension beyond an individual's birth date. This substantially limits the extent to which migration can be conclusively attributed to specific events or phenomena (e.g. conflict-induced displacement), or whether there was both the spatial and temporal overlap necessary for peer interactions to take place. Finally, many of the observed dynamics are related to specific facets of the Kurdish conflict in Turkey and may not necessarily hold in other contexts.

Nevertheless, this study presents an unprecedentedly detailed window into the relationships between urbanization, youth and insurgent recruitment. The findings herein support two overarching conclusions. The first is that the nature of urbanization is more relevant to the study of insurgency than the mere quantity thereof. Many Turkish cities have grown considerably without experiencing heightened insurgent recruitment, while others have grown modestly but acted as hubs for militants. In the Turkish context, recruitment disproportionately occurs among socially proximate individuals in urban neighbourhoods, many of whom were likely displaced by conflict. This has implications far beyond Turkey, as most conflicts generate substantial internal displacement.

The second overarching conclusion is that research on youth bulges and conflict must pay greater attention to the social and familial factors that affect young people. This study finds that factors such as peer ties, family size, and birth order may have similar effects on insurgent recruitment as they do on outcomes like educational attainment. These findings have implications for the design of social policies aimed at mitigating conflict; Stewart (2015) notes that the conflict-related employment programmes that are frequently implemented by international development and aid agencies have largely failed, but does not identify viable alternatives. As such, further research into the effect of policies geared at bolstering parental investments into their children in the context of conflict is warranted. However, the importance of political variables also suggests that a country cannot simply develop its way out of conflict using economic and social policies without addressing the political grievances of militants.

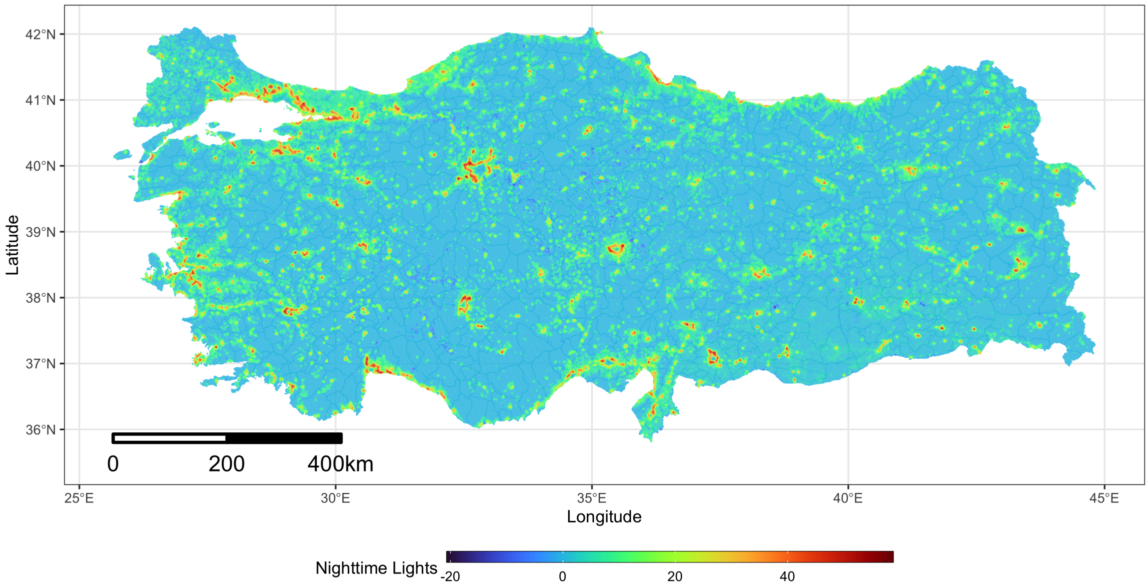
Appendix for Chapter 1

Figure A1: Underage PKK Recruits 10km Nearest Neighbours



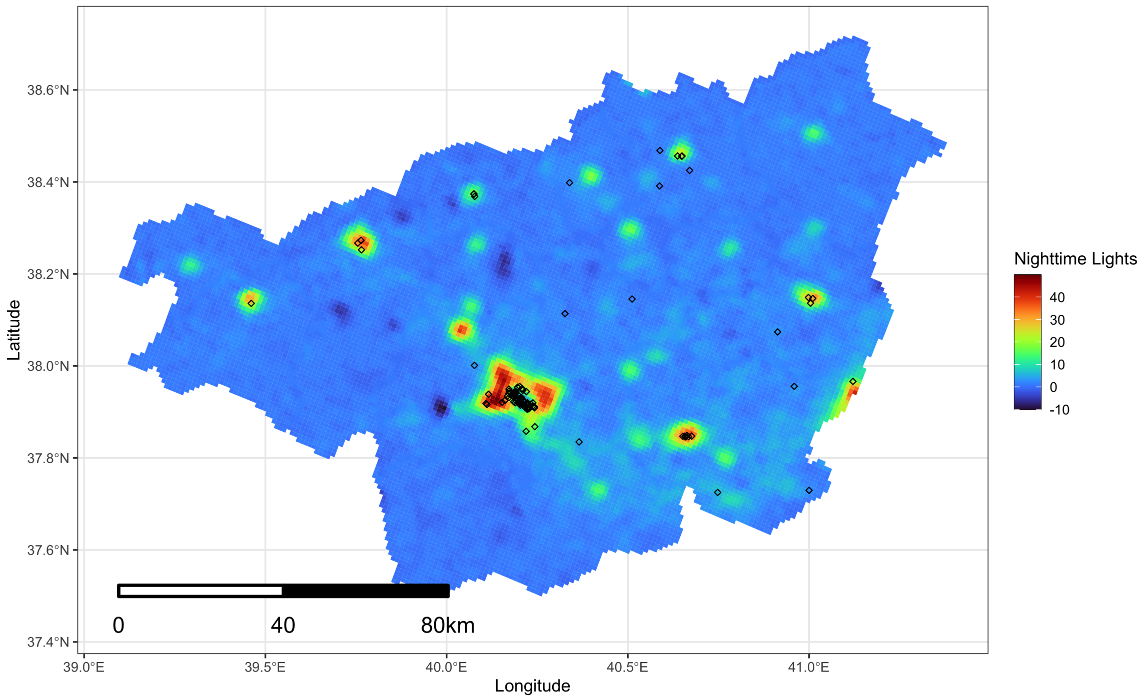
This figure displays as circles the locations of PKK members who were residents of Istanbul and joined prior to their 18th birthday. Red lines connect recruits who live within 10km of each other. Underage recruits are highly clustered within the city, virtually all living within 10km of each other.

Figure A2: Nighttime Lights Growth Across Turkey, 1992-2012



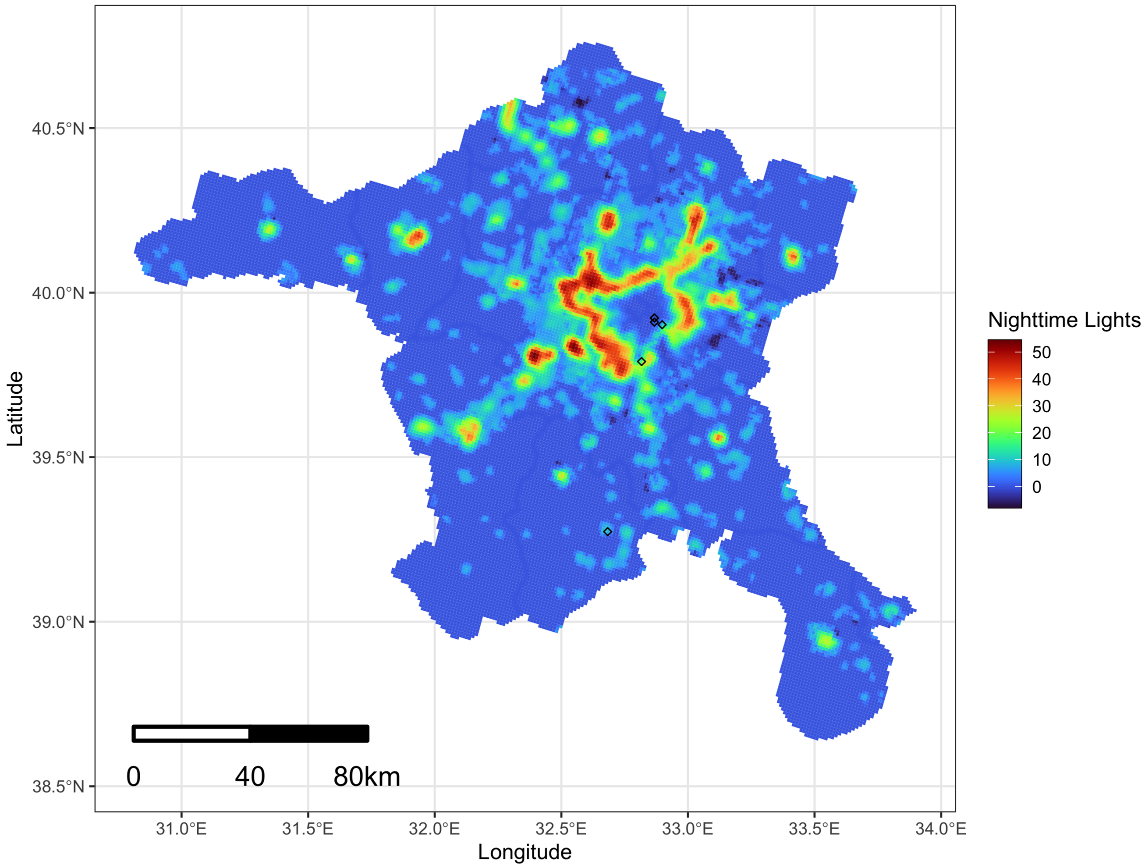
By subtracting two nighttime satellite images taken 20 years apart (one in 1992, one in 2012), this raster shows the change in nighttime luminosity across Turkey. Though changes in luminosity can be attributed to a host of phenomena including the construction of roads, methane flaring from oil and gas extraction, and industrial activity, it is most frequently associated with urban growth and rural electrification.

Figure A3: Urban Growth and Insurgent Recruitment in Diyarbakir



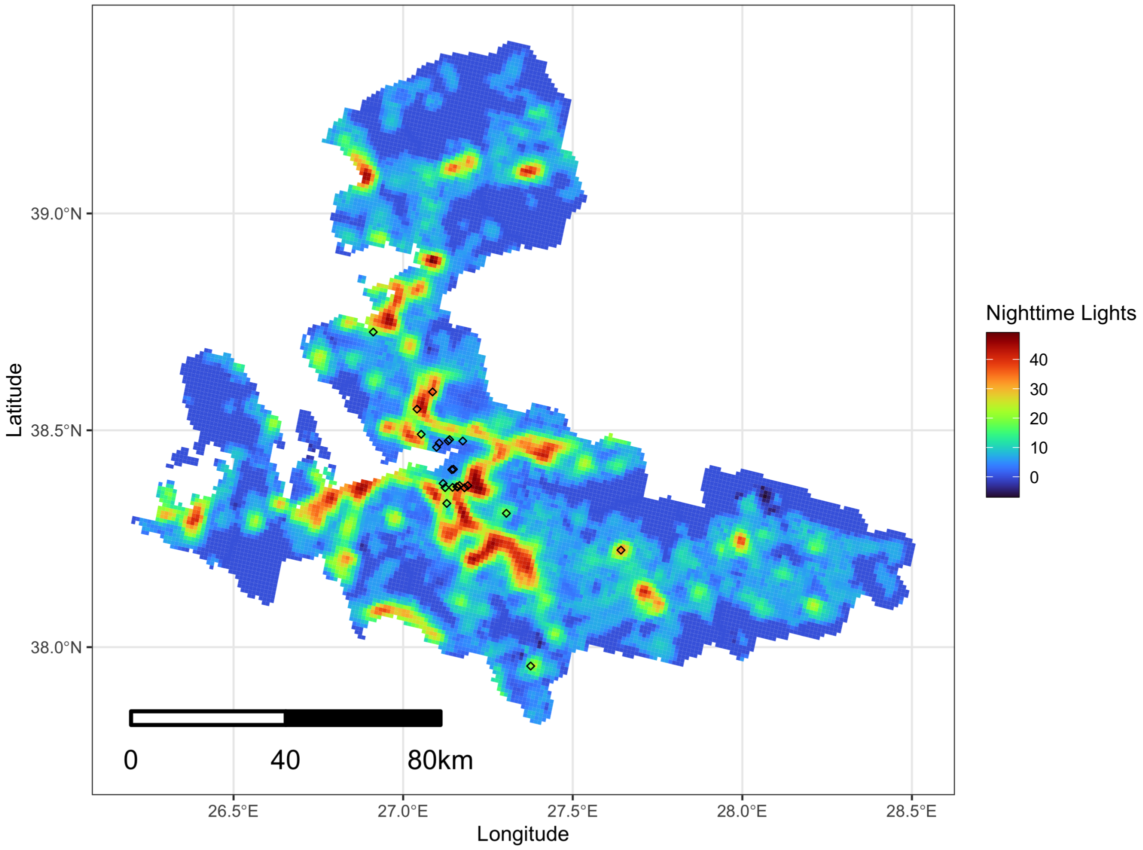
This figure displays as diamonds the locations of PKK members who were residents of Diyarbakir province prior to joining. The underlying raster shows change in nighttime lights between 1992 and 2012, with red representing growth. The city of Diyarbakir is visible near the centre of the province; a large quantity of recruits were residents of the inner city, while a smaller number lived in the urban outskirts. Outlying towns, shown as pale green dots, display both modest growth in luminosity and insurgent recruitment over the study period. Militants from rural hamlets are shown as black diamonds in blue areas. An interesting artifact in the nighttime lights data is the presence of dark spots, indicating areas that were luminous in the 1990s but have since stopped emitting light. Diyarbakir province was the site of some of the most intense fighting during the peak of the civil war in 1994, and human rights organizations have documented the burning of villages and towns by Turkish forces in this area.

Figure A4: Urban Growth and Insurgent Recruitment in Ankara



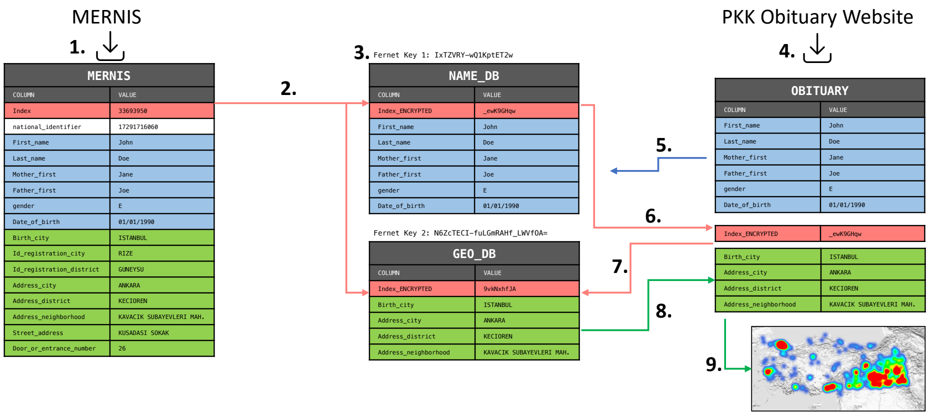
This figure displays as diamonds the locations of PKK members who were residents of Ankara prior to joining. The underlying raster shows change in nighttime lights between 1992 and 2012, with red representing growth. Though there are generally few recruits from Ankara, they live in very close proximity to each other near the city's centre.

Figure A5: Urban Growth and Insurgent Recruitment in Izmir



This figure displays as diamonds the locations of PKK members who were residents of Izmir Province prior to joining. The underlying raster shows change in nighttime lights between 1992 and 2012, with red representing growth. Izmir stands out as a favored destination for rural-urban migrants who join the PKK relative to the general population. Though the province has experienced considerable urban growth (shown in red), in keeping with the other examples, most PKK recruits originate from established neighbourhoods in the city's centre.

Figure A6: Anonymization Process for MERNIS database



1. The **MERNIS** citizenship registry is downloaded from the internet and each entry is assigned a unique index number.
2. The registry is split into two datasets:
   - **A.** **NAME_DB** contains only basic biographical information
   - **B.** **GEO_DB** contains only anonymized geographical information above the village/neighborhood level
3. The index numbers in each dataset are are encrypted using fernet keys. In this example, the index value linking the entries is 33693950. Following encryption, this becomes "ewK9GHqw" in NAME_DB and "9vkNxhfJA" in GEO_DB. Thus, the name data can no longer be linked to the geographic data without two encryption keys, stored separately.
4. Basic biographical information is collected from the PKK's public obituary website
5. Obituaries are matched with records in NAME_DB.
6. The encrypted index is extracted NAME_DB is destroyed after all obituaries are matched.
7. The indices are decrypted and matched, allowing anonymized geographic information is extracted. The encrypted index row is destroyed.
8. We are left with a fully anonymized list of approximate addresses (down to the village/neighborhood level).
9. Geographic data are geocoded, revealing the aggregate spatial distribution of recruitment

A copy of the Turkish population registry (Merkezi Nüfus İdaresi Sistemi, MERNİS) was leaked to the internet in 2012. To enable the use of this database for research purposes while maintaining respect for privacy, the following anonymization process was applied. Broadly, this involves separating the database into a table of anonymized addresses linked to a table of non-geographic demographic data. This workflow was presented to Oxford University's Central University Research Ethics Committee, which approved the use of the MERNIS database for this paper following anonymization.

Figure A7: Central University Research Ethics Committee Approval Letter

Oxford Department of International Development
Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK
Tel: +44(0)1865 281800
qeh@qeh.ox.ac.uk www.qeh.ox.ac.uk

Head of Department: Professor Diego Sánchez-Ancochea

8 July 2021

Ollie Ballinger
ODID/St Anne's College
Oxford

–

Dear Ollie

Research Ethics Approval
Ref No: SSH/ODID DREC: C1A_21_064
Title: Exploring Patterns in Kurdish Insurgent Recruitment Using Computational Methods

The above application has been considered on behalf of the Social Sciences and Humanities Interdivisional Research Ethics Committee (IDREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the Oxford Department of International Development's DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

In line with current guidance please do not undertake any data collection involving in-person interactions with participants. Once in-person research is permissible again, you will need to notify the ODID DREC or SSH IDREC via email *before* undertaking any recruitment for in-person interaction with participants. More detailed guidance is available via
https://researchsupport.admin.ox.ac.uk/governance/ethics/coronavirus.

Should there be any subsequent changes to the project that raise ethical issues not covered in the original application you should submit details to the DREC for consideration.

Yours sincerely,

Prof Loren Landau
ODID DREC Chair

Chapter 2

Assessing the impact of the Southeastern Anatolia Project on Kurdish Separatism

2.1 Introduction

The previous chapter focuses on urban recruits, but throughout it is noted that the overwhelming majority of insurgents were born in rural areas in the Southeast of Turkey. Thus, a holistic understanding of the process of insurgent recruitment must examine the conditions in the villages that militants are from, not just cities they migrate to. As previously noted, the factors that lead someone to leave behind their family and friends in search of better fortunes are likely to be intimately tied to the factors that may persuade one to take up arms against the state. Among these factors, the impact of climate shocks on civil conflict has been the subject of particular focus and debate in recent years (Koubi, 2019; Hsiang et. al., 2011; Uexkull, 2014; Couttenier and Sobeyran, 2013; Maystadt and Ecker, 2014; Miguel et. al. 2004; Miguel and Satyanath, 2011). Though many cross-country studies find a link between rainfall and civil conflict, the mechanisms underlying this relationship remain poorly understood. As a result, many of these findings have been disputed by qualitative studies that situate agricultural production within the broader context of political and social factors contributing to conflict. This is particularly true in the Middle East, where the popular thesis that a severe drought contributed to the Syrian uprising (Kelly et. al., 2015; Gleick 2014) has come under increasing scrutiny (Selby et. al., 2017; Ide, 2018). In this paper, I conduct a detailed quantitative investigation of the relationship between rainfall, irrigation, and the political economy of Kurdish separatism in Turkey.

Turkey's Southeastern Anatolia region provides an optimal setting to test both mechanisms linking agricultural income and civil conflict, as well as potential mitigating policy interventions. The region's nine provinces are characterized by an arid climate, a high degree of economic dependence on agriculture, and the highest levels of poverty in Turkey (Fikret, 2016). Situated on the border with Syria, this region experienced the same climatic shocks that have been implicated in the uprising against Bashar al-Assad, and has been the site of a Kurdish insurgency for over 40 years.

In the 1980s, the Turkish government broke ground on the Southeastern Anatolia Project (*Güneydoğu Anadolu Projesi*, henceforth GAP), a regional development program aiming to bring 1.8 million hectares of land under irrigation (GAP, 2016). Irrigation has both increased farmers' incomes by allowing them to grow higher value crops and has shielded them from ever-more frequent droughts. This, in turn, seems to strongly affect rural Kurds' views of the government; in the words of one farmer: "We did not trust the state before, but it brought electricity, water, phone, etc. to us and now we trust the state a lot" (Harris, 2009: 11).

I leverage an original, highly disaggregated dataset to investigate the relationship between irrigation and Kurdish separatism in Southeastern Turkey between 1985 and 2019. Analysis is conducted at the level of 5km-by-5km and 10km-by-10km grid cells. The primary explanatory variables include a remote-sensing based measure of irrigation, multiple drought indices, and crop production statistics. The dependent variables consist of two different measures of conflict incidence and a measure of PKK recruitment derived from online obituaries. Detailed political and economic covariates include ballot-box level election data, topographical variables, a measure of Kurdish tribal control, and nightlights. I seek to answer the following question: has irrigation reduced Kurdish separatism in Southeastern Turkey?

I first assess the cross-sectional impact of irrigation on conflict likelihood. The distribution of canal-fed irrigation schemes is largely determined by exogenous topographical constraints, as they must be below the altitude of their distributary dams and must be on nearly completely level ground. I exploit this exogenous topographical variation in the placement of irrigation schemes at the level of 5km-by-5km grid cells, during the resurgence of hostilities following the breakdown of a truce in 2015. Using a spatially autoregressive instrumental variables approach and ACLED conflict incidence data, I find that a fully irrigated 25 km^2 grid-cell is 58% less likely than the average cell to experience a conflict event involving Kurdish rebels during the 2016-2019 period. The effect is strengthened when the dependent variable is restricted to isolate military raids targeting villages that are perceived to be materially supporting the PKK. This effect is robust to the inclusion of a comprehensive array of control variables, aggressive geographic restrictions, and the inclusion of contiguous and inverse-distance spatial spillover effects.

To assess whether conflict decreases following the introduction of irrigation, I employ dynamic spatial panel models using UCDP conflict incidence data spanning over the period 1985-2018, at the level of 10km-by-10km grid-cells. Because irrigation schemes may be deliberately placed in areas that are expected to have higher potential agricultural production or less potential for conflict, the benchmark specification excludes all grid-cells that never receive irrigation during the study period. Despite the fact that the panel models and instrumental variables models use different conflict datasets and are estimated at different spatial levels, the treatment effect of irrigation is similar: a previously un-irrigated grid cell reaching the average level of irrigation in 2019 (27 km^2 experiences a 49% decrease in conflict likelihood for a given cell-year. Results remain consistent despite the use of three alternative measures of irrigation, and the use of PKK recruitment as the dependent variable rather than conflict incidence.

I present evidence on the mechanism underpinning the relationship between irrigation and conflict. Broadly, this mechanism holds that an individual's decision to join a rebellion is a function of "whether economic opportunities are so poor that the life of a rebel is attractive to 500 or 2,000 young men" (Fearon and Laitin, 2003: 88). Irrigation likely increases the opportunity cost of rebellion by generating two positive income effects: allowing farmers to grow higher value crops and insulating farmers from rainfall shocks. Yields for four major crops—which together account for over 90% of cultivated land in the region—are shown to be highly sensitive to rainfall. Wheat, which accounts for nearly 50% of the sown area in Southeastern Anatolia, is particularly sensitive. District level panel models suggests that the precarity of agricultural income plays a role in Kurdish separatism: clashes are more frequent following a poor harvest of wheat, the dominant crop in the region. Irrigation induces a shift towards the cultivation of cotton, which is both more lucrative and the only crop in the sample for which yields were uncorrelated with rainfall. I complement this quantitative analysis with a qualitative case study in Appendix B involving fieldwork interviews in an irrigation scheme during a period of drought and anti-government protest. The case study indicates that farmers who receive irrigation adopt a significantly more positive view of the government than farmers who do not.

As a further test of this mechanism, treatment heterogeneity related to land inequality is explored, showing that irrigation fails to mitigate conflict in areas with high land inequality. The prevalence of tribal social structures characterized by "feudal" land tenure means that the degree to which farmers benefit economically from irrigation is likely to be uneven across the region. Interacting the irrigation variable with various indicators of tribal control suggests that the introduction of irrigation in tribal areas actually increases the likelihood of conflict by generating violent local competition over this new source of rents. This effect is strongest in areas with a high proportion of government-aligned tribes, many of whom function as state-sponsored paramilitaries through the "Village Guard" program. Local news reports confirm that historical competition between tribes—known as *kan davalari* or "blood feuds"—are increasingly being fought over scarce irrigation water and often involve paramilitaries. This aligns closely with recent literature on rent-seeking and lootability in civil wars (Berman et. al., 2017). I further explore this phenomenon in a second qualitative case study in Appendix B. Fieldwork interviews and local news reports trace the process through which the introduction of a scarce resource in an area with a history of inter-tribal conflict led to the militarization of blood feuds and the intensification of conflict as a result of irrigation.

This paper makes three main contributions to the literature. First, I fill a critical gap in the growing literature on climate and conflict by exploiting exogenous variation in the relationship between rainfall and agricultural income to provide detailed evidence on the mechanisms linking the two. Second, I address the lack of social, historical, and political context in this literature by situating agricultural income within the broader political economy of the conflict using a detailed array of bespoke covariates. Third, I provide evidence that public infrastructure programs can partially mitigate the growing effects of climatic shocks on conflict incidence and insurgent recruitment.

The paper proceeds as follows. Section 2.2 conducts an overview of the existing scholarship on relationship between agricultural income and civil conflict. Section 2.3 provides background on the Southeastern Anatolia Project, the dynamics of PKK recruitment, and their interconnections. Section 2.4 describes the data, and Section 2.5 contains the empirical analysis. The spatially autoregressive instrumental variables approach is conducted in subsection 2.5.1, focusing on a cross-section of conflict events spanning 2016-2019. Subsection 2.5.2 examines longitudinal trends through the use of dynamic spatial panel models during the 1985-2018 period. Section 2.6 tests the "opportunity cost of rebellion" mechanism in two parts: first, by exploring the relationship between irrigation, agricultural production, and conflict; second, by examining treatment heterogeneity related to extreme land inequality in areas controlled by Kurdish tribes (*Asiretler*). Section 2.7 concludes.

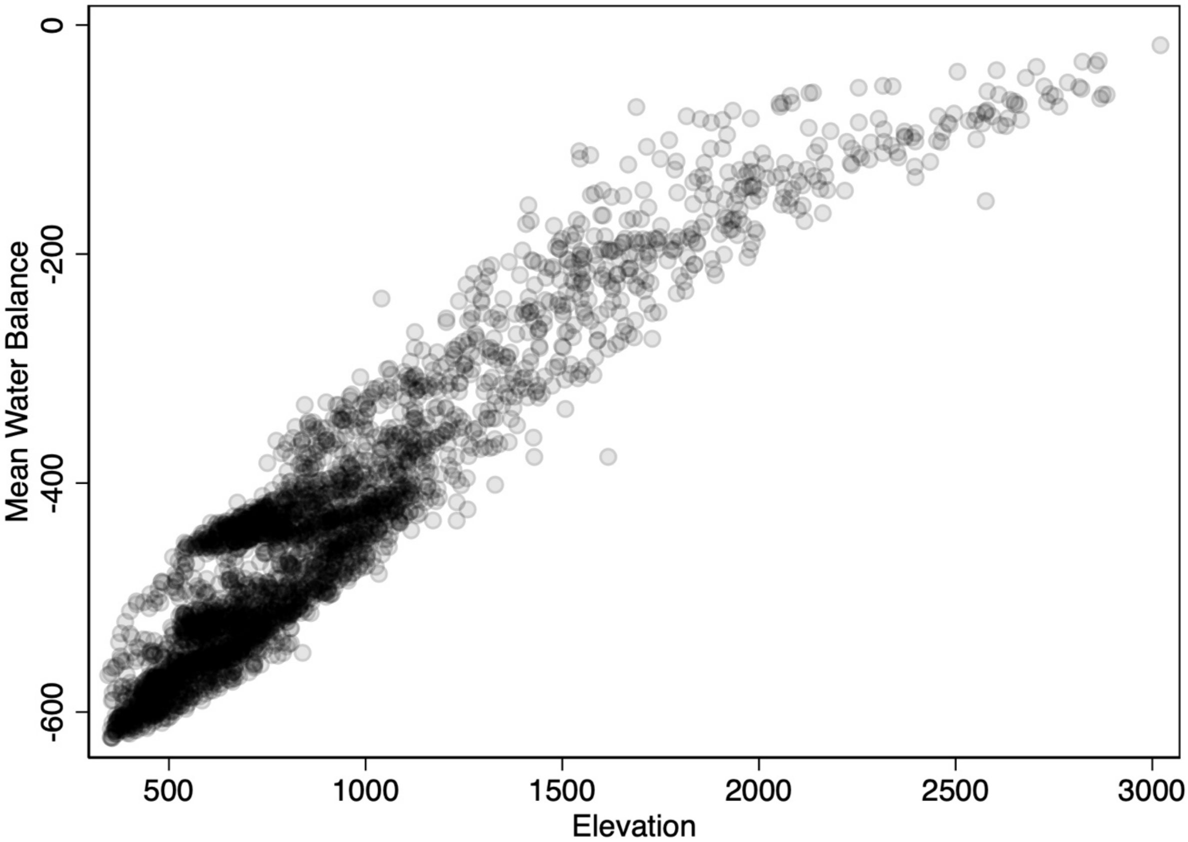
2.2 Related Literature: Development, Drought, and Rebellion.

Droughts have been argued to have a uniquely strong effect in motivating civil conflict due to their negative effect on the availability of essentials such as food and water, and because they disproportionately affect low-income rural communities engaged in labor-intensive agricultural work (McGuirk and Nunn, 2020; Hsiang et. al., 2011; Uexkull, 2014; Couttenier and Sobeyran, 2013; Maystadt and Ecker, 2014). The most well-established mechanism through which economic development affects violent conflict is through the former's effect on the opportunity cost of rebellion (Roche, Müller-Itten, and Dralle, 2020; Collier and Hoeffler, 1998). As incomes rise relative to the economic returns to participating in an insurgency, the incentive to join a rebel group diminishes (Grossman, 1991). Dal Bó and Dal Bó (2004) suggest that in a two-sector model of the economy, negative shocks to the labor-intensive (rather than the capital-intensive) sector significantly increase relative returns to joining a rebellion.

The provision of irrigation to farmers interacts with both of the mechanisms outlined above by simultaneously increasing rural incomes through crop substitution and by insulating farmers from climatic shocks. Indeed, in studies that use rainfall as an instrument for the effect of agricultural income on civil conflict, irrigation is often cited as a threat to identification, a potential policy solution, and an avenue for further research (Blattman and Miguel, 2010; Miguel et al. 2004; Bohlken and Sergenti, 2010; Harrari and LaFerrara, 2018). Despite abundant empirical evidence on the relationship between agricultural income shocks and civil conflict, the role of irrigation as a potential mitigating factor has been largely neglected. Sarsons (2015) finds that in India, income in districts downstream from dams is much less sensitive to rainfall shocks, but that this has no effect on the incidence of Hindu-Muslim riots. This is largely irrelevant to the mechanisms discussed above, however, as participation in a riot offers no economic benefits. Gatti, Baylis, and Crost (2020) provide the first evidence that irrigation attenuates the link between climate shocks and conflict in Indonesia, but acknowledge a number of identification concerns. As such, mechanisms involving the opportunity cost of rebellion and agricultural income shocks merits further investigation.

Furthermore, many of the empirical findings linking drought and civil conflict have been either weak or contested. Rainfall is frequently used as an instrument for agricultural income, and the impacts of the former on conflict are consequently treated as causal (Miguel et. al. 2004; Harrari and LaFerrara, 2018). However, rainfall variables likely violate the exclusion restriction, not least through the well-established positive relationship between rainfall and elevation (Song et al., 2019). Figure 2.1 shows that in Southeastern Anatolia, 88% of the variation in a 5km cell's mean water balance is explained solely by elevation. High levels of spatial aggregation and resulting bias from within-country heterogeneity in cropping patterns and the spatial distribution of conflict also hamper identification (Ciccone, 2011). Blattman and

Figure 2.1: Relationship Between Water Balance and Elevation



This figure displays the relationship between a 10km grid cell's mean water balance (precipitation minus evapotranspiration) and elevation in Southeastern Anatolia.

Miguel (2010: 8) contend that "the most promising avenue for new empirical research is on the subnational scale, analyzing conflict causes, conduct, and consequences at the level of armed groups, communities, and individuals." As a result, empirical analyses are increasingly being conducted using grid-cells, remote sensing data, and geospatial techniques (Harrari and LaFerrara, 2018; Verwimp et. al., 2019).

The present study addresses these theoretical and empirical gaps in the literature on the microeconomics of violent conflict. Empirical analysis is conducted at the level of 10 × 10 km and 5 × 5 km grid-cells, using two distinct measures of conflict incidence and a measure of PKK recruitment, an irrigation variable derived from remote sensing, and a rich array of georeferenced political, climatic, and economic covariates. By examining insurgent recruitment in the context of irrigation in Turkey's arid Kurdish region at such a high level of detail, fine-grained tests of the opportunity cost mechanisms of insurgent become possible. This paper makes empirical and theoretical contributions to the expanding literature on the microeconomics of violent conflict, the impact of infrastructural investments on development, and spatial econometrics.

2.3 Background

2.3.1 Kurdish Separatism

Following the establishment of the Turkish Republic in 1923, the government undertook an effort to create a nation-state modeled after European powers such as France (Park, 2013). This involved the writing of a constitution, a strong emphasis on secularism, and most importantly a policy of "Turkification", whereby "Turkish ethnic identity has been strictly imposed as a hegemonic identity in every sphere of social life" (Aktar in Ülker, 2005: 29). Later that year, a group of Kurdish tribes led by Sheikh Said rebelled against the Republican government with the goal of establishing an "independent Kurdistan" (Olson, 2000: 69). The rebellion lasted nearly four months, involved the siege of the region's largest city, Diyarbakir, and led to 8,000 casualties before Sheikh Said was defeated (Ibid).

For much of the 20th century, Southeastern Anatolia saw sporadic bouts of armed insurrection by the Kurds, followed by devastating military campaigns by the government. The development of a unified Kurdish resistance movement came in 1974, with the establishment of the Kurdistan Worker's Party (Patiya Karkerên Kurdistanê, henceforth PKK) by Abdullah Öcalan (Özcan, 2013). Thereafter, the cycles of revolt and repression were replaced with a low-intensity conflict beginning in 1984, peaking in the 1990s, and continuing to this day. Following the breakdown of a 2013-2015 truce, hostilities between the PKK and the Turkish government have intensified almost to the level of the conflict's peak in 1994. Olson (2013: 32) notes that "more Turkish military actions have been carried out in Kurdistan than in any other area of Turkish concern, foreign or domestic".

The economic destitution of Turkey's Kurdish region is an explicit feature in the writings of the PKK's founder, Abdullah Öcalan (2007; 2015). The Marxist ideological character of nearly all 20th-century Kurdish rebel movements is often attributed to the fact that “the Kurdish identity question was expressed in terms of regional economic inequalities and suggested a socialist solution” (Yavuz, 2001: 10). In a study of PKK recruitment, Tezcür (2016) finds an inverse correlation between recruitment and district-level GDP, though he makes no claims to causality. As a “peasant movement”, the PKK’s main base of support is among the rural farming villages in Southeastern Anatolia and nearby mountainous regions (Yarkin, 2015: 31).

2.3.2 The Southeastern Anatolia Project

The Southeastern Anatolia Project (*Güneydoğu Anadolu Projesi*, henceforth GAP), is a regional development program consisting of 22 dams and 19 hydroelectric power plants on the Tigris and Euphrates rivers, as well as 1.8 million hectares of irrigated land (GAP, 2016). The rate of extreme poverty in the nine provinces encompassed by the project is more than five times the national average, with 44% of the population living below the national poverty line of US$1.1/day in 2007 (Fikret, 2016; Saatci and Akpinar, 2007: 632). Nearly two thirds of all economic activity in the region is derived from agriculture, and income gains associated with the transition from rain-fed subsistence agriculture to irrigated cotton cultivation are estimated to range from three- to sevenfold (Bilgen, 2016; Tokdemir et. al., 2016: 2). Exports from Southeastern Anatolia more than quadrupled between 2005 and 2015 (GAP, 2018). There is evidence, however, that the developmental impact of the project has been highly uneven, with income gains tightly confined to irrigated areas, leaving non-irrigated areas virtually unchanged by GAP (Bakirci, 2001). Even within irrigated areas, inequality related to land tenure is significant (Fikret, 2016). Furthermore, flooding caused by the dams has displaced up to 355,000 people and destroyed cultural heritage sites such as Hasankeyf (Varsamidis, 2010). Many have been resettled into purpose-built towns nearby such as Yeni Hasankeyf (”New Hasankeyf”) less than a kilometer away from the now-flooded town). As such, externalities from this project can themselves be a powerful source of grievance.

However, ethnographic studies conducted in newly irrigated areas in Southeastern Anatolia generally found a “heightened sense of state legitimacy as a function of irrigation access"; in the words of one farmer, "our view of the state changed positively... we had hatred before, but now they started investing in the Southeast" (Harris, 2016: 9). Another farmer stated, "We did not trust the state before, but it brought electricity, water, phone, etc. to us and now we trust the state a lot" (Harris, 2009: 11). Conversely, one farmer—frustrated that his village had not received irrigation—asked, "Why should I not support Öcalan?" (Harris, 2006: 193).

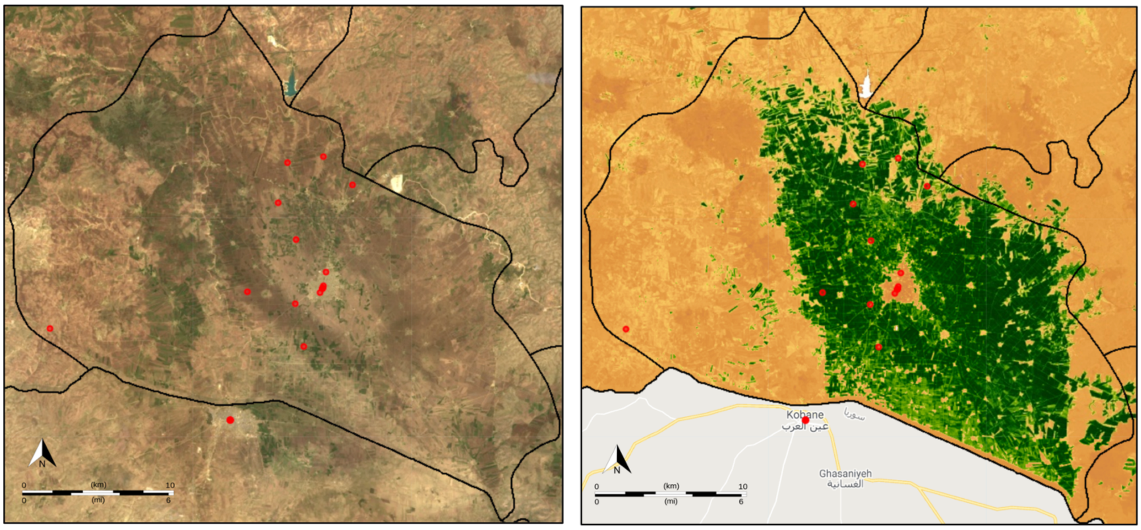
There was an explicit expectation among Turkish policymakers that "GAP will help to decrease, if not eliminate, the appeal of the PKK" (Jongerden, 2010: 137; Olson, 1996: 96). A leaked U.S. intelligence cable paraphrases then-President Abdullah Gül as saying "If Kurds are gainfully employed, have better educational opportunities, and see increased levels of infrastructure development throughout their region, their affinity for the terrorist PKK will wane further. A top priority is ensuring completion of the massive Southeastern Anatolia Project (GAP) within five years." (Wikileaks, 2008b). In a region where few farmers have enough savings to insulate themselves from income shocks, government-provided irrigation appears to have a substantial impact on their perception of the Turkish state.

2.4 Data

I constructed a dataset spanning a comprehensive array of political, economic, and climatic variables for Southeastern Anatolia. All variables are georeferenced, enabling analysis at multiple spatial scales and the use of highly disaggregated geographic units. Variables are derived from remote sensing, web scraping, and existing datasets used widely in the literature.

The primary explanatory variable in this analysis is the expansion of irrigated agricultural land, which is gauged using remote sensing. Because of overlapping growth schedules, the crop cycle in Southeastern Anatolia only supports the planting of one crop per year in a given field (Ozdogan et. al, 2002). Cotton, which accounts for 96% of irrigated crops in the region, is grown in the mid-to-late summer and harvested in early October (Unlu, 2007). Thus, if green vegetation is observed in July or August, it is almost certainly irrigated agriculture, and as such "there is overwhelming consensus

Figure 2.2: Isolating Irrigation using Summertime NDVI

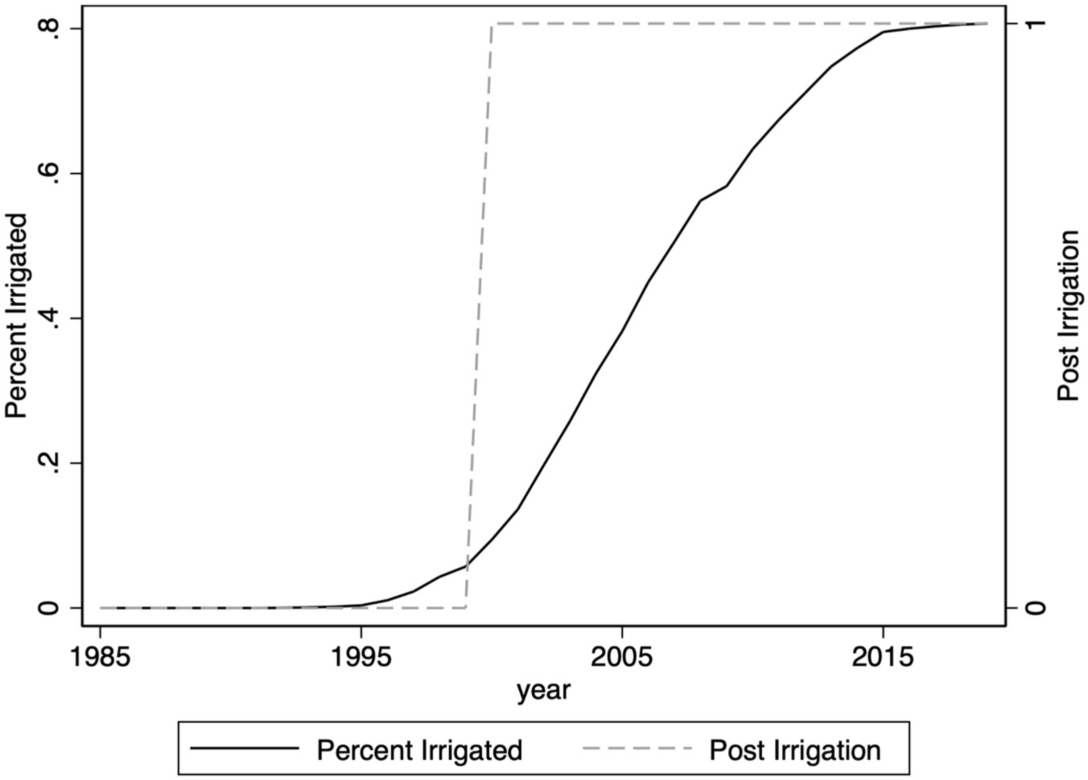


The panel on the left displays an optical satellite image of the district of Suruç, with red dots indicating the geolocated birthplaces of PKK recruits from Tezcür (2016). The panel on the right displays the Normalized Difference Vegetation Index (NDVI), calculated using Landsat 8 imagery taken during August 2020. This process clearly highlights the Suruç irrigation scheme.

on the efficiency of vegetation indices such as the NDVI in identifying irrigated fields" in Southeastern Anatolia (Ozdogan 2006). 32-day composite NDVI images were sourced for July and August from NASA's LANDSAT satellites, from 1985 to 2019. A minimum NDVI cutoff filters out wild vegetation. Figure 2.2 compares an unprocessed satellite image of the district of Suruç to a processed image, highlighting the extent of irrigation. Three distinct measures of irrigation are derived from the NDVI data. The first measure denotes the percent of a grid-cell that is under irrigation in a given year. The second is a binary variable indicating the onset of irrigation, moving from 0 to 1 when a grid-cell is more than 20% irrigated. The third denotes the number of years since irrigation was first introduced. A comparison of these three measures for a sample district is shown in Figure 2.3.

Conflict variables are derived from two main sources. Panel data on conflict were compiled from the Uppsala Conflict Data Program, one of the most widely used conflict datasets. The main conflict variable used herein is a binary measure of whether fighting occurred involving the PKK in a cell in a given year, spanning from 1989 to 2019. The frequency of clashes over time is shown in Figure B1. This includes PKK attacks on Turkish Armed Forces (TSK) and civilians, as well as TSK raids

Figure 2.3: Irrigated Area and Irrigation Onset variables

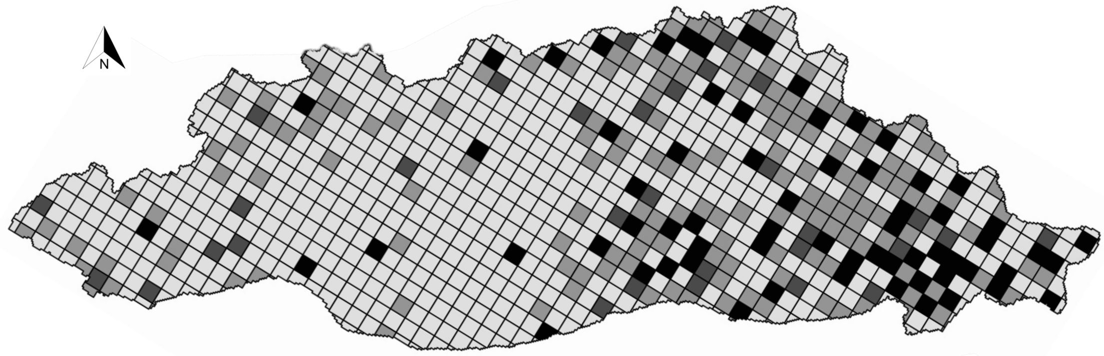


The solid line shows the percent of a 10km grid cell located in Kiziltepe that is under irrigation in a given year, derived from the remote sensing procedure carried out in Figure 2.2. The dashed line shows a binary variable indicating irrigation onset for the same cell.

on PKK positions. For added robustness, conflict data were also compiled from the Armed Conflict Location and Event Data project (ACLED). ACLED data contain more detail than the UCDP data, but only go as far back as 2016 for Turkey. As such, the ACLED data were pooled across three years (2016-2019) to create a cross-sectional variable denoting the spatial distribution of conflict events in the recent flare-up of fighting between the PKK and the government. Figure 2.4 shows the geographic distribution of conflict events from both datasets at the 10 km grid-cell level, showing a high density of clashes in the Eastern portion of the study area.

In addition to looking at conflict incidence from two different sources, data on PKK recruitment are also used to provide an even more detailed investigation of conflict channels. Data on PKK recruitment was obtained from Tezcür (2016). Each observation includes the year and location of an individual's recruitment to the

Figure 2.4: Geographic Distribution of Violent Clashes



The figure above displays the spatial distribution of violent clashes between the PKK and the Turkish Government in Southeastern Anatolia using geolocated conflict data from the UCDP and Armed Conflict Location & Event Data (ACLED). Darker cells contain more total clashes across the entire study period.

PKK. The dataset was compiled from PKK obituaries listed in two online magazines (Serxwebûn and Berxwedan) published by the PKK since 1982. Each obituary contains biographical information on the recruit including name, gender, as well as locations and dates of birth, death, and recruitment. Many even include a recruit's previous occupation, their parents' occupation, and whether they had relatives who had joined previously. There are data on 8,266 recruits in total—though the present analysis uses a subsample of 2,678 who were born and recruited in the study area—between 1985 and 2012. The dataset as a whole comprises roughly 41% of known militants based on numbers reported by both the PKK and a Turkish Parliamentary commission (Tezcür, 2016). In contrast to the vast majority of empirical analyses on conflict and development that utilize conflict incidence as a proxy for recruitment, these data contain the actual dates and geographic coordinates of recruitment, allowing for an unusually detailed investigation of individual motivations for joining a rebellion.

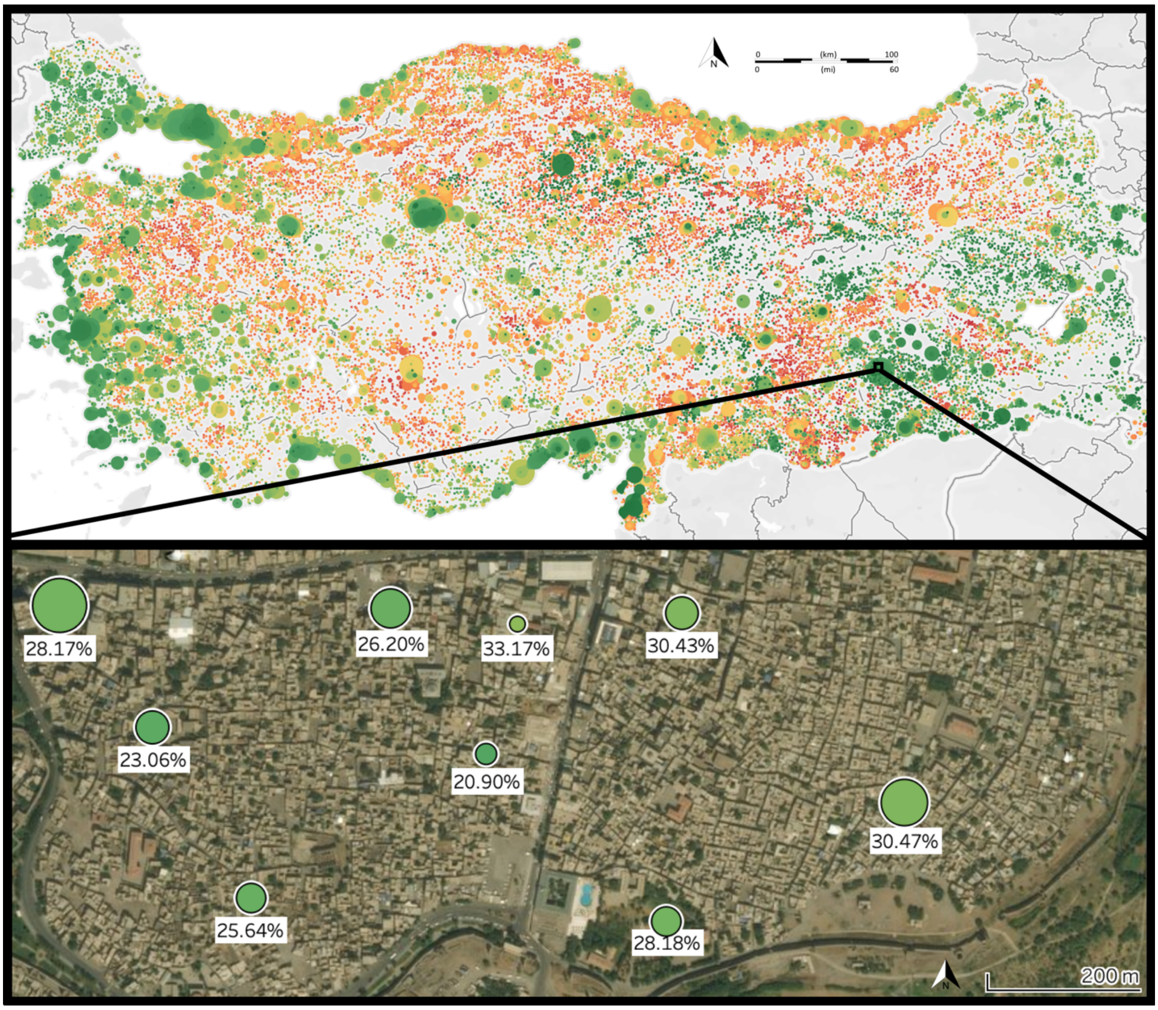
A key feature of the political economy of rural Southeasten Anatolia is the prevalence of Kurdish tribes (aşiretler). A subgroup of these tribes—mainly those that practice Sunni Islam, speak Kurmancî, and have become involved in right-wing politics—were armed by the Turkish state to form paramilitary "village guards" (köy koruçuluğu) to fight the PKK (Guida, 2014). On the other hand, a significant number of tribes—mainly those that practiced Alevi Islam, spoke Zazaki, and had become involved in socialist politics during the 20th century— have historically been at the helm of the Kurdish struggle for autonomy (Yavuz, 2001). Furthermore, accounting for aşiret control is important because land tenure is often characterized as "feudal" in tribal areas, which could temper the positive income effect of irrigation wrought by the Southeastern Anatolia Project (Öktem, 2014: 180).

The quantitative identification of tribal regions of Turkey relies on their practice of "combined voting" (birleşik oy), whereby "many villages vote en masse—usually without making a single deviance—for the party chosen by the chieftain" (Guida, 2014: 176). Scans of individual ballot box receipts are available on the Turkish Supreme Elections Board (YSK) website for all elections since 2009. Web scraping was used to access data from a total of 2.9 million ballot boxes spanning 20 elections at various levels (municipal, provincial, presidential, and parliamentary). The spatial distribution of ballot box-level election results for the whole of Turkey is shown in Figure 2.5, with a pane magnifying central Diyarbakir. Ballot boxes in which one party or candidate received over 95% of the vote were reclassified into separate dummy variables with 1 denoting bloc voting for a given party. The number of instances of bloc voting in each district or grid-cell was divided by the total number of ballot boxes therein, yielding a measure of the proportion of likely tribally administered villages in each area. 3,840 instances of birleşik oy were encountered (1564 for the HDP, 2276 for the AKP), constituting 5.36% of the total sample. This is consistent with survey findings by Akşit and Akçay (1997), who found that 4.3% of the 187 villages in their survey sample were under aşiret control.

In addition to the variables on tribal control, several additional variables were derived from ballot-box data. The respective vote shares of the AKP or HDP across all elections are measured to provide a sense of the overall political leanings of a given area. Finally, the log number of voters registered at each ballot box is used as a highly detailed proxy for population density.

climatic variables were constructed using monthly rainfall and temperature values from the University of Idaho's TerraClimate dataset. These values are used to compute the Standardized Precipitation-Evapotranspiration Index (SPEI), which unlike conventionally- used drought indices, such as the Standardized Precipitation Index (SPI) or the Palmer Drought Severity Index (PDSI), integrates not only rainfall

Figure 2.5: Ballot Box-Level Election Results



The results from 2.9 million ballot box-level election results between 2009 and 2019 are displayed above. Red indicates a high proportion of votes for the Justice and Development Party (AKP). Point size reflects the number of registered voters at a ballot box. The top panel shows results for all of Turkey, while the bottom panel shows results for several neighborhoods in Diyarbakir.

but temperature into its calculation. Harrari and La Ferrara (2018) note that while the conflict literature tends to consider the effects of either temperature or rainfall, SPEI allows for their joint effect to be estimated. The calculation of SPEI begins with the computation of the water balance, or the difference between rainfall and potential evapotranspiration. These values are standardized by cell such that the resulting value is expressed in terms of standard deviations from the cell's historical average. The main climatic variables used herein are an annual measure of SPEI, and the water balance.

To avoid neglecting the urban dimension of the conflict, nighttime lights data are used to capture cell-level information on urbanization and economic development. Nighttime lights data were collected from the Defense Meteorological Satellite Program (DMSP), and reflect the average radiance displayed by a cell in a given year. The difference in nighttime lights between 1992 and 2014 is used as a rough proxy for development at the cell level. As cities expand, highways are built, and villages receive electricity for the first time, their nighttime lights signature increases (Bruederle and Hodler, 2018). The maximum observed nightlights value in a given cell is also used to distinguish between urban, sub-urban, and rural areas.

Further time invariant cell-level covariates include elevation, slope, a binary indicator of whether a cell is on the southern border, and kilometers of roads. Finally, two longitudinal dummy variables control for different phases in the conflict; the first indicates whether or not a given year was in a period of ceasefire, and the other indicates the onset of the Syrian civil war. The former accounts for low levels of conflict resulting from armistice agreements, while the latter controls for possible cross-border spillovers from the Syrian civil war owing to the relationship between the PKK and Syrian Kurdish rebel groups such as the YPG. Summary statistics for all variables at the 5km and 10km level are reported in Table 2.1. The raw, unaggregated geospatial data for several key variables defined above are visualized in the interactive online tool.

2.5 Empirical Analysis

2.5.1 Cross-Sectional Instrumental Variables Approach

Exploiting exogenous topographic variation in the placement of irrigation schemes, this section examines the cross-sectional relationship between irrigation and conflict in Southeastern Anatolia. Employing a spatially autoregressive instrumental variables approach conducted at the level of 5km grid-cells, I show a persistent negative relationship between a cell's irrigated area and the likelihood of conflict incidence.

Using conflict incidence data from the Armed Conflict Location and Event Dataset spanning form 2016 to 2019, I first estimate a cross-sectional instrumental variables

Table 2.1: Summary Statistics

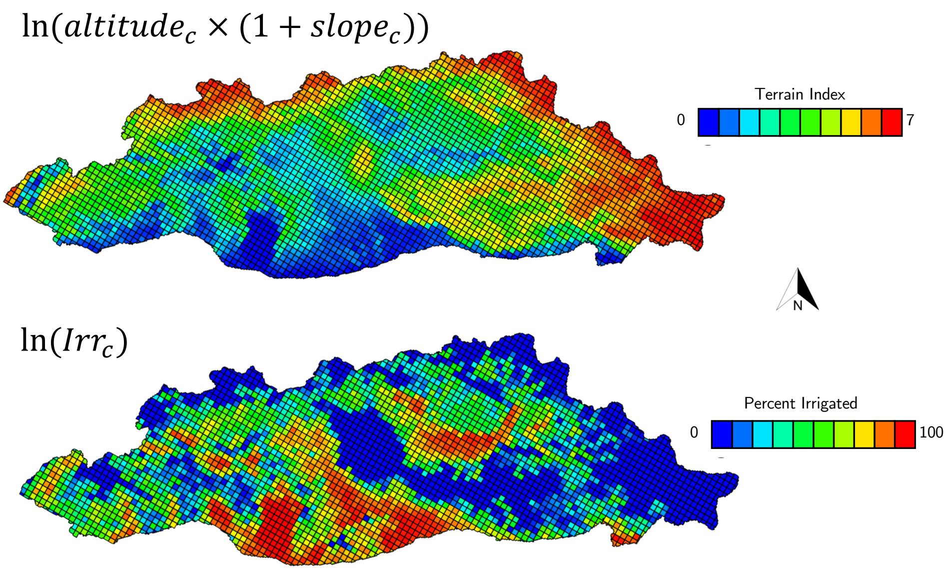
	5km Grid Cells			10km Grid Cells		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Irrigated Area	85400	0.911	3.086	27230	3.048	10.481
Post Irrigation	85400	0.154	0.361	27230	0.159	0.365
Clashes (UCDP)	73200	0.009	0.096	23340	0.033	0.179
Clashes (ACLED)	85400	0.084	0.278	27230	0.225	0.418
Population	85400	5.326	0.966	27230	5.508	0.855
AKP voteshare	85400	0.595	0.191	27230	0.588	0.164
HDP voteshare	85400	0.136	0.12	27230	0.144	0.111
Tribal	85400	0.108	0.098	27230	0.108	0.08
Nightlights Change	85400	4.658	5.072	27230	4.159	3.721
Nightlights	85400	7.88	7.301	27230	7.329	5.87
Roads	85400	9.727	1.212	27230	11.015	1.114
SPEI	82960	-0.046	0.69	26452	-0.045	0.688
Slope	85400	7.822	6.149	27230	8.536	6.595
Elevation	85400	803.996	305.47	27230	851.096	359.956
Ceasefire	85400	0.229	0.42	27230	0.229	0.42
Syrian War	85400	0.257	0.437	27230	0.257	0.437
Area	85400	24.583	2.224	27230	94.013	17.527
Border	85400	0.028	0.165	27230	0.075	0.263

This table provides summary statistics for all variables at both the 10km and 5km grid cell level.

model at the level of 5km by 5km grid-cells. Similar to Duflo and Pande (2007), who use river gradients as an instrument to assess the impact of dam construction on poverty in India, I exploit exogenous geographic variation in the placement of irrigation schemes.

Identification relies on two geographic constraints faced by gravity-fed irrigation schemes: land must be relatively low–below the altitude of distributary dams– and relatively flat to avoid runoff. Identifying irrigation schemes purely on the basis of elevation is complicated by the fact that dams in the region are built at different elevations. Similarly, the presence of flat high pastures prevents the precise identification of irrigation using only slope data. As such, the instrument used herein is defined as the log product of elevation and slope. The interaction of these terms highlights areas that are both relatively low and flat. Figure 2.6 compares the spatial distribution of the terrain instrument and the irrigation variable at the level of 5km grid-cells, showing a high degree of similarity. Equation 2.1 below specifies the general two stage

Figure 2.6: Topography and Irrigation



The top panel shows topographical variation in the study area using 5km grid cells. In the top panel, low values in this Terrain Index (blue) denote areas that are relatively low and flat, while high values (red) indicate areas that are high and steep. The bottom panel shows the distribution of irrigated cropland in red, which closely matches the distribution of low, flat terrain.

instrumental variables strategy:

$$\begin{aligned} Irr_i =& \gamma_0 + \gamma_1 Z_i + \gamma_2 X_i + \epsilon_i \\ Y_i =& \beta_0 + \pi_1 \overline{Irr}_i + \beta_1 X_i + u_i \end{aligned} \tag{2.1}$$

Where Irr_i is an endogenous variable denoting the log area under irrigation in grid cell i, Z_i is the terrain instrument, X_i is a matrix of covariates, and Y_i is the dependent variable denoting conflict incidence.

The exclusion restriction assumes that topography only affects conflict by determining geographic eligibility for irrigation. This is likely violated in the full sample: many PKK training camps and bases are located in mountainous regions that are inaccessible to Turkish ground forces (Olson, 1997). As such, there is likely to be a positive relationship between topography and conflict incidence for reasons unrelated to irrigation. However, this is only true for extremely rugged areas, as the strategic benefit of placing bases in the mountains is only incurred if the terrain effectively precludes vehicular access meaning that the exclusion restriction plausibly holds for non-mountainous areas. Mountains are clearly identifiable in right tail of the elevation histogram of Southeastern Anatolia presented in Figure B2. The solid line represents an altitude cutoff at 1200 meters that excludes mountainous grid-cells. The dotted line shows a more aggressive altitude cutoff at 1000 meters that even begins to exclude agricultural land.

Table 2.2: Irrigation and Conflict Incidence, Cross-Sectional IV Estimates

	(1) Probit	(2) Probit	(3) Probit	(4) 2SLS	(5) 2SLS	(6) 2SLS
Log Irrigated Area	-0.335***	-0.293***	-0.268***	-0.0511***	-0.0410***	-0.0373***
	(0.0410)	(0.0447)	(0.0500)	(0.00720)	(0.00696)	(0.00769)
Population	0.136***	0.153***	0.209***	0.0256***	0.0257***	0.0322***
	(0.0477)	(0.0526)	(0.0578)	(0.00775)	(0.00777)	(0.00799)
AKP voteshare	-1.271***	-1.526***	-1.092***	-0.185***	-0.207***	-0.132***
	(0.306)	(0.331)	(0.384)	(0.0477)	(0.0506)	(0.0484)
HDP voteshare	0.490	0.278	0.902	0.0516	0.00934	0.0960
	(0.465)	(0.498)	(0.562)	(0.0697)	(0.0721)	(0.0699)
Tribal	-0.0266	0.175	-0.231	0.122**	0.137**	0.0898
	(0.576)	(0.634)	(0.735)	(0.0580)	(0.0603)	(0.0618)
Nightlights Change	-0.0617***	-0.0632***	-0.0573**	-0.0174***	-0.0183***	-0.0179***
	(0.0233)	(0.0243)	(0.0231)	(0.00437)	(0.00431)	(0.00424)
Nightlights	0.0697***	0.0665***	0.0592***	0.0203***	0.0202***	0.0195***
	(0.0176)	(0.0185)	(0.0174)	(0.00321)	(0.00318)	(0.00311)
Roads	0.157***	0.217***	0.202***	0.0132***	0.0196***	0.0189***
	(0.0463)	(0.0669)	(0.0694)	(0.00320)	(0.00542)	(0.00581)
SPEI	2.698	-1.222	-5.205	0.396	-0.0959	-0.481
	(4.834)	(5.402)	(6.239)	(0.653)	(0.720)	(0.802)
Area	0.0465**	0.0328	0.0313	0.00481**	0.00489*	0.00508
	(0.0227)	(0.0257)	(0.0277)	(0.00195)	(0.00275)	(0.00321)
Previous Clashes	0.00375**	0.00337*	0.00234	0.00155*	0.00131	0.000911
	(0.00184)	(0.00179)	(0.00163)	(0.000839)	(0.000818)	(0.000783)
Border	0.707**	0.709**	0.694**	0.0690	0.0782	0.0805
	(0.280)	(0.294)	(0.311)	(0.0441)	(0.0502)	(0.0537)
Observations	2,440	2,223	1,965	2,440	2,223	1,965
Geographic Restriction	None	<1200	<1000	None	<1200	<1000

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

This table reports IV estimates of the effect of irrigation on PKK recruitment at the level of 5km by 5km grid cells, with robust standard errors displayed in parentheses. Equation 2.1 is estimated as a two-stage probit model in columns 1-3 of Table 2, and as a linear probability model via 2SLS in columns 4-6. Columns 1 and 4 use the full sample, columns 2 and 5 restrict the sample to observations below 1200 meters, and columns 3 and 6 further restrict the elevation cutoff to 1000 meters. The Kleibergen-Paap F statistic is 1697 for the full sample and 1743 for the maximum geographic restriction.

Equation 2.1 is estimated as a two-stage probit model in columns 1-3 of Table 2.2, and as a linear probability model via 2SLS in columns 4-6. Columns 1 and 4 use the full sample, columns 2 and 5 restrict the sample to observations below 1200 meters, and columns 3 and 6 further restrict the elevation cutoff to 1000 meters. Table 2.2 indicates a consistent negative relationship between irrigation and conflict incidence in the 2016-2019 period. This effect persists across model specifications and geographic restrictions. The magnitude, direction, and significance of the coefficients is not substantively different between the probit model and the linear probability model. The Kleibergen-Paap F statistic is 1697 for the full sample and 1743 for the maximum geographic restriction, suggesting that the instrument is not weak.

However, due to the 5km resolution of the grid, spatial autocorrelation must be accounted for. The structure of spatial dependence is modeled using a spatial weighting matrix W. The benchmark W is a binary contiguity matrix which assigns a value of 1 to the eight cells directly bordering the cell of interest, and 0 to all others. The spatial lag of a variable X for cell c is denoted as $X_c \times W$ and is equal to the average value of X in the eight cells neighbouring c. For robustness, a spatial weighting matrix denoting the inverse distance from cell c is also used. Figure B3 shows the effect of spatially lagging the irrigation variable using contiguity and inverse-distance weighting matrices. The contiguity matrix models spillovers in the direct vicinity of irrigation schemes, while the inverse distance matrix allows for gradually attenuating spatial dependence across the entire study region. Equation 2.2 augments the instrumental variables model by adding three spatially autoregressive terms:

$$\begin{aligned} Y_i =& \beta_0 + \pi_1 \overline{Irr}_i + \beta_1 X_i + \theta_1 WX_i + \lambda_1 WY_i + u_i \\ u =& \rho W u + \epsilon_i \end{aligned} \tag{2.2}$$

WY_i is a spatial lag of the dependent variable, WX_i is a spatial lag of the independent variables, and u is a spatially autoregressive error term. The result is a Spatially Autoregressive Instrumental Variables (SAR-IV) model.

Columns 1-4 in Table 2.3 report the results of equation 2.2 where W is a binary contiguity spatial weights matrix, while columns 5-8 use an inverse distance matrix. All columns employ the most extreme geographic restriction (<1000 meters). The dependent variable in columns 1 and 5 is general conflict incidence. Subsequent columns disaggregate conflict into its components: the incidence of PKK attacks (columns 2 and 6), military raids (columns 3 and 7), and protest incidence (columns 4 and 8). The same set of time invariant cell-level control variables used in the panel models is included but values are calculated at the 5km level. Two additional control variables are specified: historical water stress is measured as the cell-level mean of the rainfall variable, and temporal persistence of conflict is accounted for by taking the sum of pre-2000 clashes using the UCDP conflict data.

Table 2.3: Irrigation and Conflict Incidence by Type, Spatially Autoregressive IV Estimates

	W= Binary Contiguity Matrix				W= Inverse Distance Matrix			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All Conflict	PKK Attacks	Military Raids	Protests	All Conflict	PKK Attacks	Military Raids	Protests
Log Irrigated Area	-0.0134**	-0.00646*	-0.0132***	-0.00349	-0.0192**	-0.0180***	-0.0199***	-0.00563
	(0.00549)	(0.00362)	(0.00452)	(0.00352)	(0.00774)	(0.00601)	(0.00612)	(0.00391)
Y x W	0.497***	0.372**	0.466**	0.145	1.605	-3.066	0.976	3.480***
	(0.162)	(0.157)	(0.195)	(0.177)	(1.178)	(2.085)	(1.162)	(1.316)
Population	0.0453***	0.0306***	0.0349***	0.0216***	0.0400***	0.0298***	0.0312***	0.0207***
	(0.00898)	(0.00681)	(0.00791)	(0.00590)	(0.00818)	(0.00629)	(0.00709)	(0.00556)
AKP voteshare	-0.0514	0.0350	-0.0823	0.0336	-0.0872	0.0195	-0.113**	0.0120
	(0.0588)	(0.0371)	(0.0526)	(0.0248)	(0.0542)	(0.0353)	(0.0471)	(0.0237)
HDP voteshare	0.121	0.121*	0.0908	0.0840*	0.0717	0.0539	0.0367	0.0255
	(0.116)	(0.0660)	(0.110)	(0.0510)	(0.110)	(0.0636)	(0.0993)	(0.0483)
Tribal	0.0186	-0.0525	0.0724	-0.0303	0.0328	-0.0571	0.0869	-0.0489
	(0.0691)	(0.0418)	(0.0568)	(0.0401)	(0.0729)	(0.0451)	(0.0606)	(0.0425)
Nightlights Change	-0.0247***	-0.0256***	-0.0263***	-0.0268***	-0.0198***	-0.0221***	-0.0224***	-0.0240***
	(0.00382)	(0.00363)	(0.00364)	(0.00290)	(0.00380)	(0.00346)	(0.00360)	(0.00304)
Nightlights	0.0307***	0.0289***	0.0279***	0.0300***	0.0241***	0.0240***	0.0234***	0.0250***
	(0.00297)	(0.00290)	(0.00299)	(0.00260)	(0.00268)	(0.00249)	(0.00268)	(0.00238)
Roads	0.0224***	0.0152***	0.0134***	0.0134***	0.0218***	0.0128***	0.0127**	0.00894***
	(0.00613)	(0.00458)	(0.00495)	(0.00388)	(0.00644)	(0.00463)	(0.00503)	(0.00336)
SPEI	6.753***	5.313***	5.849***	2.135	6.605***	6.853***	4.463***	2.344**
	(2.195)	(1.678)	(1.827)	(1.344)	(2.186)	(1.913)	(1.418)	(1.191)
Area	0.00851**	0.00529*	0.00527*	0.00347	0.00573*	0.00311	0.00334	0.000598
	(0.00356)	(0.00274)	(0.00317)	(0.00251)	(0.00330)	(0.00257)	(0.00298)	(0.00236)
Previous Clashes	0.000663	0.00101	0.000851	0.000906	0.000808	0.00103	0.000970	0.00105
	(0.000561)	(0.000752)	(0.000671)	(0.000697)	(0.000602)	(0.000777)	(0.000699)	(0.000766)
Border	0.0284	0.00257	0.0220	-0.000650	0.0527	0.0261	0.0213	0.0125
	(0.0577)	(0.0424)	(0.0473)	(0.0379)	(0.0552)	(0.0399)	(0.0433)	(0.0345)
Observations	1,965	1,965	1,965	1,965	1,965	1,965	1,965	1,965
Geographic Restriction	<1000m	<1000m	<1000m	<1000m	<1000m	<1000m	<1000m	<1000m

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

This table reports spatially autoregressive IV estimates of the effect of irrigation on PKK recruitment at the level of 5km by 5km grid cells, with robust standard errors displayed in parentheses. Columns 1-4 report the results of equation 2.2 where W is a binary contiguity spatial weighting matrix. Columns 5-8 use an inverse distance matrix. All columns employ the most extreme geographic restriction (1000 meters). Spatial lags of the independent variables are included in the models, but coefficients are not reported. The dependent variable is general conflict incidence in columns 1 and 5, PKK attacks in columns 2 and 6, military raids in columns 3 and 7, and protest incidence in columns 4 and 8. The same set of time invariant cell-level control variables used in the panel models is included but values are calculated at the 5km level.

The results are instructive. Table 2.3 displays a consistent negative relationship between irrigation and conflict incidence regardless of whether spatial dependence is assumed to be confined to neighbouring cells, or to decrease linearly with distance. Though the overall effect of irrigation across all types of violent conflict is negative, there is heterogeneity in the strength and significance thereof.

Columns 1 and 5 indicate that irrigation strongly decreases the likelihood of experiencing a conflict event. Column 1 indicates that a fully irrigated 25 km^2 grid-cell is 58% less likely than the average cell to experience a conflict event involving Kurdish rebels during the 2016-2019 period. Disaggregating conflict shows that this effect is primarily driven by a negative relationship between irrigation and military raids. These "counter-terror operations" are typically carried out in villages that are believed to be materially supporting the PKK by harboring militants, serving as weapons caches, or producing improvised explosive devices. Columns 2 and 6 suggest that the relationship between irrigation and the likelihood of experiencing a PKK attack is much weaker.

The locations of military raids are a better proxy for popular support of the Kurdish insurgency than the locations of PKK attacks; the safehouses and weapons caches targeted by these raids are unlikely to be placed in villages where guerillas do not feel they have the trust of the local population. The link between popular support for the PKK and the latter's choice of targets is less clear. On the one hand, opportunistic attacks against local security forces or police officers would likely be higher in areas with more militants. On the other hand, rebels probably do not carry out offensive attacks against the villages that support them—if anything, the opposite is likely to be true. Indeed, column 7 shows that military raids are less likely to occur in areas that consistently vote for the ruling AKP, while column 8 shows an insignificant but positive relationship between AKP vote share and the likelihood of experiencing a PKK attack. Thus, if irrigation erodes popular support for the PKK, this would be most visible through the former's effect on military raids. To illustrate this point, Table B1 contrasts the qualitative descriptions of sample PKK attacks and military raids drawn from the ACLED dataset.

The coefficients of the control variables further contextualize these results. There are strong indications that the resurgence of conflict in the lowlands of Southeastern Anatolia is largely taking place in inhabited areas rather than isolated hideouts and bases. Population, nighttime luminosity, and road cover are positively correlated with all types of conflict across all models. However, despite the strong positive relationship between overall luminosity and conflict, areas that experienced the greatest growth in nightlights are significantly less likely to experience conflict of any kind. Nightlights growth is strongest in rural villages that have received electricity for the first time (much of which is derived from hydropower), peri-urban areas on the outskirts of growing cities, and areas that have received substantial infrastructural investments such as roads and highways. This trend provides further contextual evidence that areas that have benefitted most from state-led development policies display lower levels of anti-government militancy. The fact that irrigation remains significant despite controlling for general economic development suggests that the latter's influence on conflict is not purely economic. It also further rules out omitted variable bias related to the potential endogeneity of infrastructural investments by accounting for uneven development across the region.

Conflict continues to display spatial and temporal dependence. The number of violent clashes experienced by a grid cell before the first ceasefire in 2000 is a consistently strong predictor of conflict in the current period. The spatially lagged dependent variable provides an insight into the nature of conflict spillovers in the region. Columns 1-3 indicate that the likelihood of experiencing conflict in a given cell is substantially increased if any of the neighbouring cells have experienced a conflict event. This is true for all types of conflict. When an inverse distance matrix is employed WY becomes insignificant, suggesting that conflict spillovers are highly localized.

To summarize, exploiting exogenous geographic variation in the placement of irrigation schemes, this section finds that the resurgence of violence in Southeastern Anatolia between 2016 and 2019 is systematically lower in irrigated areas. This effect is robust to the inclusion of a wide array of control variables, aggressive geographic restrictions, and the inclusion of contiguous and inverse distance spatial spillover effects. Irrigation most strongly affects the likelihood of experiencing a military raid, which, in turn, acts as a proxy for popular support of the PKK; this is because raids are typically carried out against villages that are believed to be materially supporting the rebels.

2.5.2 Dynamic Spatial Panel Models

I next examine the longitudinal relationship between irrigation and Kurdish separatist conflict in Southeastern Turkey, using two sets of panel models. The first is estimated via OLS with standard errors corrected for spatial and serial correlation (Hsiang, 2010), and the second is a spatial Durbin model. The guiding question in this instance is whether separatism decreases in a given area following the introduction of irrigation. Analysis is conducted at the level of 10km by 10km grid cells. At higher resolutions, conflict events become too rare to be estimated consistently in the panel models. Conversely, decreasing the resolution of the cells hampers the precise delineation between irrigated and non-irrigated areas. The main dependent variable is a binary indicator of PKK-involved conflict incidence derived from the UCDP dataset. The primary explanatory variable is a binary measure indicating the onset of irrigation.

For a panel of N cells and t years, Irr denotes the irrigation variable, X a vector of time-invariant controls, and L a vector of time-varying controls. Fixed effects for year and district$d$ are represented by γ and μ, respectively, and a linear time trend is denoted by τ:

$$Y_{c,d,t} = \beta_0 + \beta_1 Irr_{c,t} + \beta_2 X_c + \beta_3 L_t + \gamma_t + \mu_d \tau_c + \epsilon_{c,d,t} \tag{2.3}$$

A linear probability model is estimated via OLS, using panel-adjusted Conley (1999) standard errors developed by Hsiang (2010), thereby accounting for both spatial and serial correlation.

Duflo and Pande (2007: 3) outline several concerns related to the endogeneity of infrastructural investments, specifically related to dam construction and irrigation. Key among these is the fact that the provinces that are chosen for dam construction projects are often targeted for political and economic reasons including potential agricultural productivity or the political clout of local governments.

As such, all analysis is restricted to observations within the nine provinces that were selected for the Southeastern Anatolia Project. While the selection of these provinces was very likely endogenous, the first stage IV results from the previous section suggest that the distribution of gravity-fed irrigation schemes within this area was largely determined by topographical factors. The use of a 10km grid further allays these concerns; though it is plausible that even the districts within these nine provinces were also endogenously selected for treatment, the feasibility of such targeting (both administratively and practically) at the level of 10km grid cells is low. Nevertheless, the benchmark specification excludes all grid-cells that never receive irrigation. As such, only cells that eventually become irrigated but have yet to receive irrigation are treated as counterfactuals for irrigated cells.

Several complexities in the dynamics of conflict remain to be accounted for. Harrari and La Ferrara (2019) note that the spatial dependence of conflict can be disaggregated into two main categories: autocorrelation in conflict determinants, and direct spillovers from fighting into neighbouring areas. Furthermore, they expect conflict to exhibit strong serial correlation.

To model spatial and temporal dependence, a dynamic, spatially autoregressive Durbin model (SDM) is estimated using maximum likelihood.

$$\begin{aligned} Y_{c,d,t} =& \beta_0 + \beta_1 Irr_{c,t} + \beta_2 X_c + \beta_3 L_t + Y_{c,d,t-1} \\ &+ \rho W + \theta_1 Irr_{c,t} \times W + \theta_2 X_c \times W + \gamma_t + \mu_c \tau_c + \epsilon_{c,d,t} \end{aligned} \tag{2.4}$$

The temporal persistence of conflict is integrated into the model with the addition of a one-year lag of the dependent variable. The addition of spatially lagged independent variables accounts for autocorrelation in conflict determinants, while the spatial lag of the dependent variable captures variation in conflict incidence owing to direct contagion.

The selection of fixed effects is informed by the inclusion of a comprehensive array of time-invariant cell-level characteristics including elevation, slope, distance to the border, population, nighttime lights, historical party vote shares, tribal control, and road cover. Combined with the already narrow geographic focus of the study area, these variables either directly or indirectly control for many of the innate political, economic, ethnic, and geographic differences between cells. As such, the main threat posed by omitted variables relates to time-varying trends in the political and -economic history of the region. These include elections at all levels, which likely affected the intensity of the government's approach to both infrastructural investment and Kurdish separatism. Year fixed effects (γ_t) control for national-level shocks. Another is forced displacement: many areas experienced long-term outmigration as a result of climate shocks, conflict, or poverty. District- and cell-level linear time trends ($\mu_d \tau_d$ and $\mu_c \tau_c$) control for these unobserved phenomena.

Table 2.4 reports the results from equations 2.3 and 2.4. Columns 1 and 2 employ the full sample of observations, while columns 3 and 4 restrict the sample to only include areas that receive irrigation at some point during the study period. Columns 1 and 3 are estimated using a panel OLS model with standard errors corrected for serial and spatial autocorrelation (Hsiang, 2010). Results in columns 2 and 4 correspond to spatial Durbin Models estimated via Maximum Likelihood, using a binary contiguity spatial weighting matrix.

As the results show, there is a strong negative relationship between the onset of irrigation and the likelihood of experiencing conflict in a given cell-year across all models. Results derived from the full sample of observations suggest that the introduction of irrigation reduces conflict likelihood by between 1.2% and 1.9%. Though this effect may appear small, it corresponds to between 37% and 58% of the unconditional mean of the dependent variable. However, results derived from the full sample are potentially driven by endogeneity in the selection of irrigation schemes. This is partially mitigated through the inclusion of a fixed effect for the "treatment" group which denotes that a cell received irrigation at some point during the study period, accounting for innate differences between areas that were selected for irrigation and areas that were not. For added robustness, columns 3 and 4 omit all cells that never receive irrigation. The negative effect of irrigation on conflict incidence not only survives the sample restriction, it is slightly strengthened.

Results are also robust to the inclusion of an array of control variables, fixed effects, and model specifications. The dynamic spatial Durbin models include a one year lag of the dependent variable, which shows that conflict in the previous year

Table 2.4: Irrigation and Conflict Incidence, Spatial Panel Models

	(1)	(2)	(3)	(4)
	OLS HAC I	Durbin I	OLS HAC II	Durbin II
Post Irrigation	-0.0193***	-0.0124***	-0.0189***	-0.0224***
	(0.00675)	(0.00445)	(0.00666)	(0.00428)
Treatment Group	-0.00144	0.00150		
	(0.00955)	(0.00487)		
Y_{t-1}		0.407***		0.249***
		(0.0199)		(0.0461)
Population	0.0376***	0.0133***	0.00359	0.000133
	(0.00814)	(0.00274)	(0.0105)	(0.00330)
AKP voteshare	0.0148	-0.0278**	-0.0299	-0.00312
	(0.0425)	(0.0137)	(0.0435)	(0.0155)
HDP voteshare	0.289***	0.0167	-0.0558	0.0746*
	(0.0553)	(0.0255)	(0.0689)	(0.0395)
Tribal	-0.0456	-0.0398**	0.0860	0.0973***
	(0.0547)	(0.0198)	(0.0595)	(0.0337)
Nightlights Change	0.00125	0.00178	0.00412	0.00166
	(0.00546)	(0.00136)	(0.0127)	(0.00600)
Nightlights	0.00332	0.00157*	0.0115	0.00696*
	(0.00338)	(0.000827)	(0.00799)	(0.00399)
Roads	0.0144***	0.0102***	-0.00694	-0.00495
	(0.00454)	(0.00143)	(0.00636)	(0.00320)
SPEI	0.00699	-0.00664	0.00419	0.0208*
	(0.00648)	(0.0156)	(0.00999)	(0.0117)
Ceasefire	-0.212***	-0.0740***	0.00154	-0.0128
	(0.0308)	(0.0197)	(0.0154)	(0.0120)
Syrian War	0.270***	0.0472	0.00482	0.0244
	(0.0418)	(0.0306)	(0.0123)	(0.0177)
Border	0.00409	-0.0172**	-0.0231	-0.0167***
	(0.0192)	(0.00686)	(0.0206)	(0.00618)
Observations	23,340	22,562	6,510	6,293
Year FE	X	X	X	X
Cell Specific Time Trend	X		X	
District Specific Time Trend		X		X
Only Treated Cells			X	X

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

This table reports estimates of the effect of irrigation on PKK recruitment from spatial panel models at the level of 10km by 10km grid cells, with robust standard errors displayed in parentheses. Columns 1 and 2 employ the full sample of observations, while columns 3 and 4 restrict the sample to only include areas that receive irrigation at some point during the study period. Columns 1 and 3 are estimated using a panel OLS model with standard errors corrected for serial and spatial autocorrelation (Hsiang, 2010). Columns 2 and 4 are spatial Durbin Models estimated via Maximum Likelihood, using a binary contiguity spatial weighting matrix. All specifications include year fixed effects. Columns 1 and 3 include a cell fixed effect interacted with a linear time trend. Columns 2 and 4 include a district fixed effect interacted with a linear time trend. Slope, elevation, and cell area are included as variables but coefficients are not shown.

significantly increases the likelihood of experiencing conflict in the following year. Restricting the sample to only include irrigated areas does not substantially affect the association between irrigation and conflict, the coefficients of several control variables are affected. Covariates associated with urban land cover are particularly affected; in the full sample, these is a strong positive association between conflict and both population and road cover. When the sample is restricted to irrigated cells, both of these relationships become insignificant. This likely results from the omission of most large cities in columns 3 and 4. Interestingly, the ceasefire variable follows the same pattern, suggesting that the 2001-2013 truce saw a decrease in urban clashes, but not in rural fighting. Finally, restricting the sample to irrigated areas has a strong effect on the association between conflict and tribal voting behaviour. In the full sample Durbin model (column 2) there is a significant negative relationship between the two, while in the restricted sample Durbin model (column 4), a strong positive relationship emerges. This result is explored in detail in section 2.6.

Results from the panel models are robust to a number of alternative specifications of both the dependent and independent variables. This section explores various modifications of the benchmark model presented in column 4 of Table 2.4 above. This baseline specification is a spatial Durbin model restricted to irrigated areas, employing a binary contiguity weighting matrix, year fixed effects and a district-level linear time trend, a full set of covariates, and a one-year lag of the dependent variable. Table B2 in the appendix reports the results of these alternative specifications.

Alternative measure of conflict. The dependent variable in columns 1 and 2 is the UCDP conflict incidence measure. However, it is substituted in columns 3 and 4 for an obituary-derived measure of PKK recruitment developed by Tezcür (2016). This measure is also binary and denotes the year and location of an individual's recruitment to the PKK. As previously explored, clashes (and particularly military raids) primarily take place in villages that are perceived to be materially supporting the PKK. Measuring recruitment is a more direct way of assessing an area's support for the PKK, though there are potential biases posed by the fact that the measure is derived from PKK-reported obituaries.

Alternative measures of irrigation. Instead of the binary irrigation onset variable, two additional measures of irrigation are employed. Columns 2 and 4 use irrigated area as the treatment variable, which represents the number of square kilometers of the grid cell under irrigation in a given year. By comparing grid cells in different irrigation schemes over time, Figure 2.3 shows that irrigated area varies both cross-sectionally, as some grid cells become more irrigated than others, and longitudinally as irrigation is phased in over time. Both of these sources of variation allow for an examination of whether treatment intensity affects conflict.

The effect of irrigated area on conflict incidence is reported in column 2. In 2019, the average irrigated grid cell contained 27 km^2 of irrigated cropland, representing roughly a quarter of its total area. Though the marginal effect of each additional square kilometer of irrigated cropland on the likelihood of conflict incidence is small, the cumulative effect is considerable. A 27 km^2 increase in irrigated area leads to a 0.91% decrease in the likelihood of conflict incidence. Given that the average likelihood of experiencing a conflict event in a given cell-year is 1.8%, this effect is equivalent to nearly 50% of the unconditional mean of the dependent variable. In other words, receiving the average quantity of irrigation over the whole study period decreases the likelihood of experiencing a conflict event by roughly half.

Column 4 reports the effect of irrigated area on PKK recruitment, which yields similar results. A 27 km^2 increase in irrigated area leads to a 0.62% decrease in the likelihood of PKK recruitment, equivalent to 39% of the latter's unconditional mean. A cell in the top quartile of irrigated area (receiving 38 km^2 of irrigation over the whole period) would experience a 56% decrease in the relative likelihood of experiencing a recruitment event, and a 69% decrease in the likelihood of conflict incidence. These results suggest that greater treatment intensity leads to greater reductions in the likelihood of experiencing conflict and recruitment events.

The treatment variable in columns 1 and 3 is a measure of the number of years that have elapsed since the introduction of irrigation. Unlike irrigated area, this variable does not vary cross-sectionally and isolates longitudinal treatment intensity. Though there is a significant effect on both conflict incidence and recruitment, the effect on the latter is smaller. This might reflect the different time periods covered by each model; the recruitment data spans from 1985-2012, while the UCDP data covers 1989-2018.

These alternative specifications also shed light on the relationship between conflict and some of the covariates. There is a strong negative relationship between historical electoral support for the AKP and PKK recruitment, and a consistent positive relationship between HDP support and conflict incidence. Tribal voting behaviour is strongly correlated to both conflict incidence and PKK recruitment.

These results indicate a consistent negative relationship between irrigation and Kurdish separatism in Southeastern Turkey which is robust to the inclusion of a highly detailed array of control variables, fixed effects, time trends, and spatial and temporal dependence structures add further validity. This relationship persists whether irrigation is measured as a binary phenomenon with a before and after period, as the number of square kilometers under irrigation in a cell-year, or as the number of years since its onset. It also persists whether Kurdish separatism is measured as conflict incidence or PKK recruitment.

2.6 Mechanisms

Both cross-sectional and panel models indicate a persistent negative relationship between irrigation and conflict incidence in Southeastern Anatolia. This section examines the possible mechanisms underlying this relationship using mixed methods.

I first explore the relationship between irrigation, agricultural production, and conflict via district-level fixed-effect panel models. I find that irrigation allows crop yields to be independent from rainfall, and that there is a positive relationship between conflict and the cultivation of rainfall-sensitive wheat.

Next, I examine heterogeneity in the treatment effect related to tribalism. I find that irrigation actually increases the likelihood of conflict incidence in tribal areas due to violent competition between tribes over scarce irrigation water. The negative effect of irrigation on conflict not only persists, but is strengthened when accounting for heterogenous effects in tribal areas.

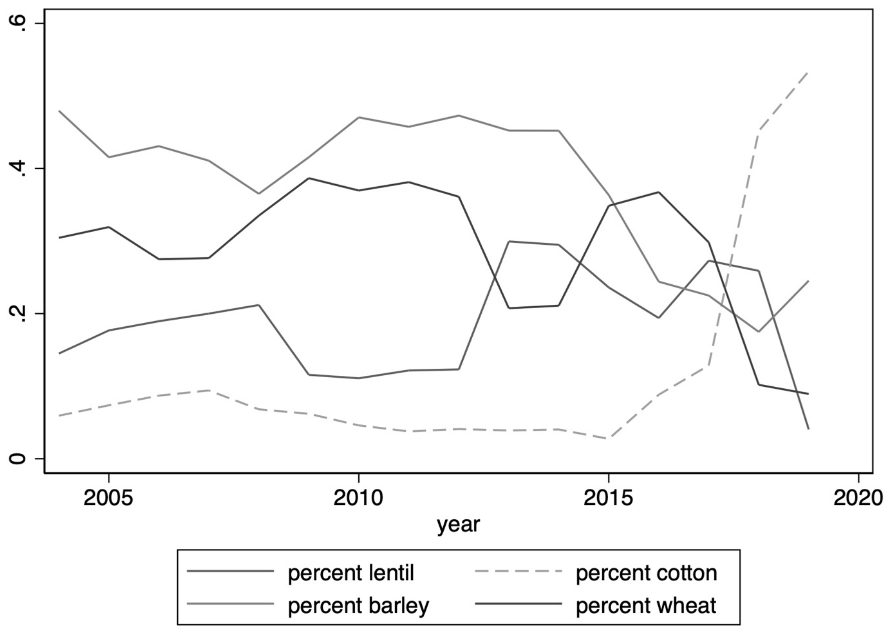
Finally, to generate a more detailed understanding of these mechanisms, I conduct two qualitative cast studies that draw on original fieldwork, ethnographic evidence, and local news reports to examine two outliers in the relationship between irrigation and conflict: Suruç and Batman. These case studies are presented in the appendix to this chapter. The first finds that despite very strong anti-government sentiment in the urban center of Suruç, interviews with farmers in the surrounding area suggest a heightened sense of state legitimacy as a function of access to irrigation. The second explores the mechanics of violent competition over resources in a tribal area with high levels of irrigation.

2.6.1 Agricultural Income

The dominant mechanism linking agricultural income and conflict in the literature is the opportunity cost of rebellion (Collier and Hoeffler, 1998). Low agricultural income levels and negative income shocks decrease the opportunity cost of joining an insurgency (Blattman and Miguel, 2010). Writing before the onset of large-scale irrigation, White (1998: 148) notes that "most of the economically active Kurds in the Kurdish region's agricultural sector are very poor—often destitute—sharecroppers. In virtually none of these cases do these Kurds have adequate savings to insulate them from further adversity during harder economic times". As such, the opportunity cost of joining the PKK is likely to be very low for rainfall-dependent farmers, and perhaps even negative in times of severe drought. Irrigation likely increases the opportunity cost of rebellion by generating two positive income effects: allowing farmers to grow higher value crops and insulating farmers from rainfall shocks. This section examines the relationship between irrigation, cropping patterns, and conflict in greater detail using district-level crop production data.

Figure 2.7 shows the composition of the crop mix of Southeastern Anatolia between 2004 and 2019. The introduction of irrigation is often followed by a drastic shift in the crop mix in favor of higher-value cotton. Five crops—wheat, barley, cotton, lentils, and hazelnuts—account for over 95% of the cultivated land in the region. Prior to irrigation, wheat is by far the dominant crop, making up over half of the total sown area in most years. Figure 2.7 plots the relative proportions of crops

Figure 2.7: The Effect of Irrigation on the Crop Mix in Suruç



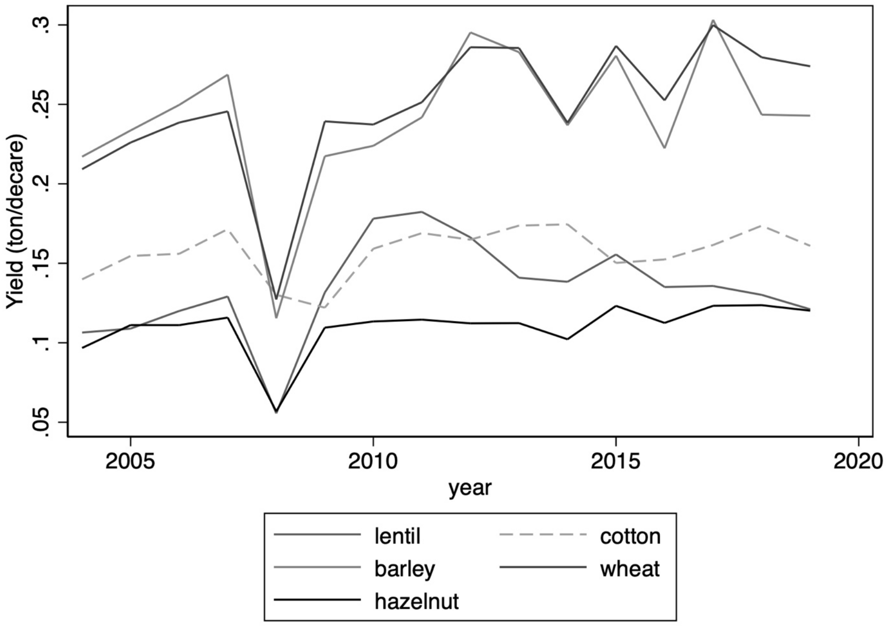
Prior to the introduction of irrigation in 2016, the dominant crop in Suruç was barley. Lentils and wheat were also grown interchangeably. Following the introduction of irrigation, the production for all three of these crops declines sharply, while production of cotton surges.

grown in the district of Suruç over time. The primary crops were barley and wheat until the onset of irrigation in 2015, which led to a dramatic shift towards cotton production. In 2019, the price per kilogram of cotton (3.2 TL) was nearly triple that of barley and wheat (1.2 TL), despite relatively similar yields (TUIK, 2020). Thus, the transition to irrigated agriculture likely generates significant income gains through crop substitution.

Irrigation in Southeastern Anatolia also effectively insulates farmers from income shocks associated with rainfall deficits. In 2008, Southeastern Anatolia suffered a severe drought, with the average water balance 24% below the mean. The impact thereof on crop yields is visible in Figure 2.8. The yields for all crops other than cotton were between 54% and 61% lower than average. Yet despite having the highest water requirement out of the main crops grown in the region, cotton saw the smallest decline in yields.

To investigate this formally, I estimate district-level fixed effect panel models estimating the effect of rainfall and irrigation on crop yields in Southeastern Anatolia

Figure 2.8: Crop Yield Sensitivity to Rainfall Shocks



The yields for five major crops over time are shown above. In 2008, Southeastern Anatolia experienced a severe drought. Yields for all non-irrigated crops decline sharply. Cotton, which is almost exclusively irrigated, suffered a much smaller decline in yields.

Table 2.5: Irrigation and Crop Yields at the District Level

	(1) Cotton Yield	(4) Wheat Yield	(2) Barley Yield	(3) Lentil Yield	(5) Hazelnut Yield
Irrigated Area	0.331***	0.651***	0.305**	0.0775	0.0231
	(0.0665)	(0.0972)	(0.143)	(0.123)	(0.160)
Water Balance	0.000878	0.0637***	0.0730***	0.0236**	0.0299***
	(0.00598)	(0.0180)	(0.0259)	(0.0115)	(0.0108)
Observations	926	974	974	955	946
Year FE	X	X	X	X	X
District FE	X	X	X	X	X

$***p < 0.01, **p < 0.05, *p < 0.1$
This table reports the results of district-level fixed effect panel models estimating the effect of rainfall and irrigation on crop yields in Southeastern Anatolia between 2004 and 2018. All specifications include district and year fixed effects.

between 2004 and 2018. The results are reported in Table 2.5. There is a strong positive relationship between rainfall and yields for all crops except cotton, confirming the trend observable in Figure 2.8. Yields for barley, lentils, and hazelnuts are unaffected by the proportion of a district under irrigation. However, a 10% increase in district-level irrigation is associated with a 64kg/decare increase in wheat production and a 31kg/decare increase in cotton production. The fact that wheat is both highly sensitive to rainfall but also more productive in districts with irrigation suggests that the majority of wheat is rain-fed, but a small proportion thereof is irrigated. If the proportion of irrigated wheat were high, we would expect yields to be largely independent of rainfall, as is the case for cotton.

In Table 2.6 I examine the relationship between crop production statistics and conflict at the district level. The production statistics consist of a one-year lag of the yield and sown area of a given crop in a given district.

There is a significant negative relationship between wheat yields and conflict incidence, suggesting that the risk of conflict is higher following a poor wheat harvest. There is also a significant positive relationship between the sown area of wheat and conflict incidence, suggesting that districts that are highly reliant on this rainfall-sensitive crop are also more conflict-prone. Both of these results are highly consistent with the mechanism that agricultural income shocks decrease the opportunity cost of rebellion. Though we might expect a negative relationship between yields and conflict for all rain-fed crops, the effect is only observable for wheat because it accounts for over 50% of the cultivated land in the region.

As previously mentioned, an important secondary effect of increasing agricultural incomes through a state-led development program involves reshaping rural populations' conception of the state, which has deep implications for insurgency. A qualitative case study in Appendix B provides a more extreme test of the capacity of irrigation to change individuals' perception of the government's legitimacy in the context of an active insurgency. In the villages across the Suruç Plain, 69 individuals joined the PKK before the introduction of irrigation. Only six joined after. Even though crop failures resulted from the state's mismanagement of irrigation infrastructure in the summer of 2018, farmers mostly acknowledged temporary hardships, but largely dismissed them in reference to the memories of drought and famine in the time before the canals were dug. Keeping in mind the positive association between crop failures and violent conflict, it appears as though the material security afforded by this state

Table 2.6: Conflict Incidence and Lagged Crop Production

	(1)
	Clashes
Ceasefire	-0.284***
	(0.0474)
Syrian War	0.122***
	(0.0288)
Cotton Yield$_{t-1}$	0.236
	(0.245)
Barley Yield$_{t-1}$	0.0171
	(0.180)
Hazelnut Yield$_{t-1}$	0.449
	(0.399)
Wheat Yield$_{t-1}$	-0.496**
	(0.235)
Lentil Yield$_{t-1}$	0.170
	(0.289)
Cotton Sown Area$_{t-1}$	-0.145
	(0.142)
Barley Sown Area$_{t-1}$	-0.0682
	(0.0939)
Hazelnut Sown Area$_{t-1}$	-0.112
	(0.472)
Wheat Sown Area$_{t-1}$	0.0356
	(0.0287)
Lentil Sown Area$_{t-1}$	-0.341**
	(0.163)
Observations	10,745
R-squared	0.122
Year FE	X
District FE	X

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

This table reports the results of a district-level fixed effect panel model estimating the relationship between crop yields and conflict incidence in Southeastern Anatolia between 2004 and 2018. All specifications include district and year fixed effects. Units for the sown area variables are presented in '000 hectares.

project insulates farmers from the hardships and grievances that drive individuals to join the PKK.

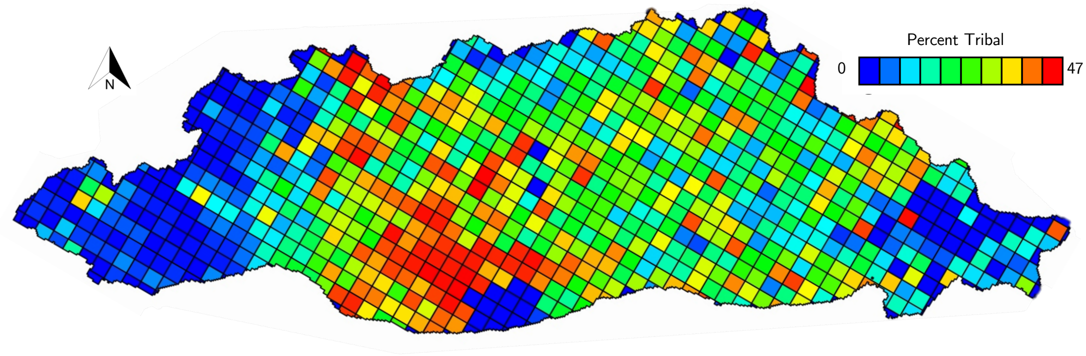
In the context of extreme poverty, low savings rates, and widespread subsistence agriculture, the opportunity cost of joining the PKK is likely to be very low particularly during times of economic hardship. The relationship between conflict and agricultural production suggests that economic precarity plays a role in Kurdish separatism: clashes are more frequent in districts reliant on rain-sensitive wheat, particularly following a poor wheat harvest. The introduction of irrigation is often accompanied by a shift in production to higher value crops such as cotton, and the decoupling of crop yields from rainfall. In conjunction, these dynamics suggest that the effect of irrigation on agricultural production and income is an important mediating factor in the relationship between irrigation and Kurdish separatism.

2.6.2 Tribal Treatment Heterogeneity

The previous section has identified the incomes generated by irrigation as one of the likely mechanisms linking irrigation and conflict. However, the distribution of these incomes is unlikely to be homogenous across the study area, particularly given the prevalence of tribal social structures which have been described as “feudal” (Aksit and Akcay, 1997). This section explores heterogeneity in the treatment effect related to tribal social structures.

The Turkish (and before it, Ottoman) government has relied on tribal chieftains to rule Southeastern Anatolia; these local elites were allowed to maintain power and collect rents from their territory on the condition that they align themselves with the government (Guida, 2014: 178). The most significant recent iteration of this phenomenon was the formation of the “Village Guard” [koy koruculugu] system in 1985, through which the government selectively armed loyalist tribes as a paramilitary force to fight the PKK (ibid). At the height of the conflict there were 95,000 village guards, equivalent to nearly half the number of total active military personnel in the entire Turkish Armed Forces (Oktem, 2014; TSK, 2014). The spatial distribution of ballot boxes displaying tribal bloc voting is shown in Figure 2.9.

Figure 2.9: The Spatial Distribution of Tribes in Southeastern Anatolia



The choropleth map above indicates the proportion of ballot boxes in a 10km cell displaying tribal bloc-voting behaviour.

Thus, an alternative explanation for the reduction in Kurdish separatism following the introduction of irrigation could be the government's use of irrigation to co-opt tribal chieftains. A stronger treatment effect in tribal areas would support this mechanism. On the other hand, a diminished treatment effect in tribal areas would strengthen the opportunity cost of rebellion mechanism; because the distribution of land in tribal areas is highly unequal, income gains associated with irrigation are unlikely to directly benefit farmers.

To explore treatment heterogeneity related to tribalism, the benchmark spatial Durbin Model (Table 2.4, column 4) is estimated with the addition of interactions between the irrigation variable and indicators of tribal voting behaviour. This allows the effect of irrigation on conflict to vary depending on whether it is introduced to land controlled by tribes AKP- or HDP-affiliated tribes. The results of these specifications are reported in Table 2.7. Column 1 adds an interaction between the treatment variable and the total proportion of ballot boxes displaying tribal bloc voting in a given grid-cell to the benchmark SDM. The coefficient on this interaction is both highly significant and positive, suggesting that conflict likelihood actually increases when irrigation is introduced into tribal areas. Columns 2 and 3 disaggregate the tribal indicator, isolating bloc voting for the AKP and HDP, thereby accounting for the political alignment of tribes.

In columns 3 and 4, the total effect of irrigation on conflict incidence becomes positive. The irrigation variable remains negative and significant across all specifications, but the interaction term is positive and significant when AKP-aligned tribes are isolated in column 3. This trend persists in column 4 when both disaggregated interaction terms are included. In column 3, the onset of irrigation reduces the likelihood of conflict incidence by 4.4% in non-tribal areas. However, in areas where AKP-affiliated tribal control exceeds 46%, the effect of introducing irrigation becomes positive. This level of tribal control is rare– only 7% of cells in the sample are above this threshold. These counter-intuitive results highlight an important facet of the political economy of violence in Southeastern Anatolia: tribal competition over scarce resources.

The effect of selective government support and co-optation of Kurdish tribes, especially during the civil war, was the militarization of "blood feuds" [kan davalari] between tribes. Ozcan (2006: 5) notes that "the tribal phenomenon helps to explain the sense of disunity among Kurds, the absence of accord, and the pitilessness of internal clashes". Akpolat (2009: 423) conducted fieldwork on one such feud which left 44 men, women, and children dead in (heavily irrigated) Mardin province, and found that the cause of these feuds is "the struggle over ownership of the primary means of production in this area: land, livestock, and water". This struggle over the means of production occurs in what has been described as an "anarchic" environment, where the citizenry—particularly the paramilitary Village Guards— is more heavily armed than the Jandarma and the local police (Criss, 1995). Because irrigation schemes sometimes cover hundreds of thousands of hectares, they often encompass the territories of several tribes. Thus, when irrigation water is rotated from villages in one tribe's territory to another, this too can become a potential site of conflict.

Local news reports have covered "blood feuds" over access to irrigation in Silvan, Sirnak, Sanliurfa, Kisehir and beyond (Hurriyet, 2007; Karar, 2019). Mahmut Tezcan (1981: 64), an anthropologist studying these feuds at Anakara University, provides a

Table 2.7: Tribal Treatment Heterogeneity

	(1) All Tribes	(2) HDP Tribes	(3) AKP Tribes	(4) AKP and HDP Tribes
Post Irrigation	-0.0498***	-0.0253***	-0.0441***	-0.0508***
	(0.00920)	(0.00533)	(0.00792)	(0.00977)
Post x Tribal	0.217***			
	(0.0500)			
Post x HDP Tribal		0.101		0.161*
		(0.0847)		(0.0896)
Post x AKP Tribal			0.0978***	0.106***
			(0.0222)	(0.0241)
Y x W	0.245***	0.249***	0.245***	0.245***
	(0.0462)	(0.0461)	(0.0461)	(0.0462)
Population	0.00167	0.000321	0.00129	0.00170
	(0.00336)	(0.00325)	(0.00336)	(0.00333)
AKP voteshare	-0.0158	0.00111	-0.0188	-0.0135
	(0.0162)	(0.0156)	(0.0164)	(0.0161)
HDP voteshare	0.0699*	0.0660	0.0803**	0.0671
	(0.0400)	(0.0431)	(0.0400)	(0.0433)
Tribal	-0.0157	0.0888**	0.00386	-0.0178
	(0.0455)	(0.0346)	(0.0425)	(0.0463)
Nightlights Change	0.000868	0.00188	0.000776	0.00104
	(0.00598)	(0.00600)	(0.00599)	(0.00598)
Nightlights	0.00739*	0.00685*	0.00744*	0.00731*
	(0.00398)	(0.00398)	(0.00398)	(0.00397)
Roads	-0.00548*	-0.00503	-0.00530*	-0.00546*
	(0.00313)	(0.00322)	(0.00317)	(0.00315)
SPEI	0.0204*	0.0205*	0.0207*	0.0201*
	(0.0117)	(0.0117)	(0.0117)	(0.0118)
Slope	0.00124	0.00157	0.00127	0.00129
	(0.00169)	(0.00170)	(0.00169)	(0.00171)
Elevation	6.52e-05	4.36e-05	6.23e-05	6.46e-05
	(4.10e-05)	(3.98e-05)	(4.07e-05)	(4.11e-05)
Ceasefire	-0.0117	-0.0124	-0.0120	-0.0114
	(0.0119)	(0.0120)	(0.0119)	(0.0119)
Syrian War	0.0262	0.0245	0.0258	0.0261
	(0.0177)	(0.0177)	(0.0178)	(0.0177)
Cell Area	-4.99e-05	-0.000123	-4.28e-05	-5.78e-05
	(9.19e-05)	(9.47e-05)	(9.48e-05)	(9.53e-05)
Border	-0.0141**	-0.0169***	-0.0141**	-0.0143**
	(0.00621)	(0.00622)	(0.00623)	(0.00627)
Observations	6,293	6,293	6,293	6,293
Year FE	X	X	X	X
District Time Trend	X	X	X	X

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

Results in this table represent modifications of the specification found in column 4 of Table 2.4. The dependent variable in all cases in the UCDP-derived conflict incidence measure. Column 1 adds an interaction between the treatment variable and the total proportion of ballot boxes displaying tribal bloc voting in a given grid-cell. Columns 2 and 3 interact the treatment variable with a disaggregated tribal indicator, isolating bloc voting for the AKP and HDP respectively. Column 4 includes both interaction terms.

detailed account of an irrigation-induced blood feud in Kirsehir that decimated several generations from two clans:

> "Cuhadar Gedik's father used to fight with Necip Aslan about water. 'You are going to irrigate now? No, I am going to irrigate now!' and so on. They fight, and Cuhadar's father shoots four of Necip's relatives and a bystander. Then one of their fathers shoots back and injures him. He shoots the brother of the man whose father attacked him. Then his father shoots back. Cuhadar's grandmother is hit in the crossfire and dies. [...] Cuhadar goes to some relatives who he saw shooting guns at a wedding in 1966 and borrows a gun. Then he goes to get coffee. Necip's brother is also there. They fight and pull guns on each other. Mehmet Aslan gets shot in the arm. He dies a while afterwards. Cuhadar is 14 years old. He goes to jail for 19 years. While he is in jail, Cuhadar's aunt is killed."

This type of local cleavage between tribes resulting from competition over irrigation water was exploited by the government and turned into a new front in the battle against the PKK. One article in a national newspaper reads:

"In the village of Otlubayir in Agri province, Muhtar [village head] Vehbi Aksoy and his relatives began a feud with Nihat Aslan, a Village Guard, over the issue of irrigation water. Village Guard Aslan fired on the Muhtar and his relatives with an automatic Kalashnikov rifle. Muhtar Aksoy and his nephews Serkan, Sinan, Hakan, and Vahit were killed, and three others were injured". (Hurriyet, 2010)

The Kalashnikov in question was given to Aslan by the government to aid in the fight against the PKK but it was used in a feud over irrigation water. The guerrillas also frequently get involved in these feuds, as they did in Şirnak, when a bus carrying 12 Village Guards to the construction site of an irrigation scheme was ambushed by the PKK, killing all passengers (CNN Turk, 2018).

A parliamentary inquiry into the Village Guard system found this to be a systematic phenomenon: "Several Village Guards have killed villagers with whom they have blood feuds based on the pretext that the latter are PKK supporters" (Belge, 2011: 107). Between 1985 and 1997, 17 village guards were prosecuted for participating in blood feuds (ibid). The militarization of these feuds was also perpetuated by the PKK itself—McDowall (2003: 421) notes that "Rather than assaulting the agha class as a whole, the PKK operated with fine calculation, exploiting blood feuds where these existed, helping to create them where they did not".

The evidence above suggests that competition over access to irrigation can spark or intensify blood feuds that have abounded in the Southeast. Kalyvas (2003: 487) contends that "framing civil wars in binary terms is misleading" and that "local cleavages and intracommunity dynamics must be incorporated in theories of civil war". The divisions between tribes that often lead to these feuds have been fostered by governments since the Ottoman era, with the most recent iteration of this policy involving the selective armament of pro-government tribes to fight the PKK through the Village Guard system. The result, as demonstrated, is that irrigation can actually intensify conflict where competition over resources overlaps with these historically reinforced divisions.

A qualitative case study in Appendix B focuses on conflict dynamics in the Batman-Silvan Irrigation Scheme, an area containing the two grid-cells in the sample that contain the highest levels of both irrigation and tribal presence. A combination of fieldwork interviews, prior ethnographic work, and local news reports details the processes through which the introduction of a lucrative economic resource in an area with pre-existing social and political cleavages can lead to violence. In this case, many of the economic benefits that were shown in the previous subsection to have a negative effect on conflict are siphoned off by tribal landlords. Furthermore, historical blood feuds between tribes intensified following the introduction of a scarce and valuable irrigation water, and these gradually became militarized as the government selectively armed tribes loyal to the government to act as a vanguard against the PKK.

An increase in conflict resulting from the introduction of a new high-value resource aligns closely with the literature on rent-seeking and lootability in civil wars. Recently, Berman et. al. (2017) found that roughly a quarter of violence in Africa involved control over mining facilities, and that clashes increased when mineral prices were high. They motivate their study by analyzing a feud between two tribes in Darfur over control of a gold mine that left 800 dead. In the context of Southeastern Anatolia, tribal control mitigates the negative effect of irrigation on conflict. In areas with exceptionally high levels of AKP-affiliated tribal control, the introduction of irrigation appears to generate violent local competition over this new source of rents. This result further strengthens the opportunity cost mechanism discussed in the previous section as tribal areas tend to exhibit extreme land inequality, negating the positive income effect of irrigation for farmers.

2.7 Conclusion

This paper provides a highly detailed test of the dominant mechanism linking agricultural income and civil conflict. This has direct bearing on the large and ever-growing body of literature examining the relationship between rainfall, income, and civil war. A fully irrigated 25 km^2 grid-cell was found to be 58% less likely than the average cell to experience a conflict event involving Kurdish rebels during the 2016-2019 period. Between 1985 and 2019, a 27 km^2 increase in irrigated area decreased the likelihood of experiencing a conflict event in a 100 km^2 grid-cell in a given year decreased by 49% relative to the mean. Agricultural income is likely the mediating factor: conflict incidence is higher in districts reliant on rain-sensitive wheat, particularly following a bad harvest. Irrigation triggers a shift from rain-sensitive crops to cotton, which is both more lucrative and resistant to rainfall shortages. Accounting for heterogeneity in the relationship between irrigation and conflict related to unequal land tenure in tribal areas strengthens the negative effect of irrigation on Kurdish separatism, providing further evidence for the opportunity cost mechanism. However, the onset of irrigation in tribal areas is associated with an increase in conflict due to militarized competition between tribes over irrigation-derived rents.

These findings have policy relevance both within and beyond Turkey. Strong serial dependence in conflict incidence, as well as the negative association between recruitment and irrigation, suggest that a military approach to the Kurdish insurgency is counterproductive. Instead, insulating farmers from climatic shocks and enhancing their livelihoods through irrigation appears to more effective in quelling separatism. The extent to which this is true in other contexts is subject to several caveats. The

PKK was established as a "peasant movement", which recruited heavily in agricultural communities. There was direct ethnographic evidence that irrigation mediated Kurdish farmers' affinity for either the Turkish state or the PKK (Harris, 2006; 2009; 2016). There was also evidence that the Turkish state explicitly expected GAP to reduce the appeal of the PKK (Wikileaks, 2008a; 2008b). Thus, climatic vulnerability, irrigation, and insurgent recruitment were directly linked in the political economy of Southeastern Anatolia, in ways that may not be true in other cases.

Further research is needed to establish the validity of the conclusions drawn herein in other contexts at a comparable level of spatial and temporal disaggregation. A more detailed empirical inquest into the concrete mechanics of the "opportunity cost of rebellion" mechanism would also be beneficial. For example, given that Dal Bó and Dal Bó (2004) propose that negative income shocks would have opposite effects in capital-intensive versus labor-intensive sectors, more work is needed to account for the fact that agriculture tends to become significantly more capital-intensive over time.

Appendix for Chapter 2

Table B1: Qualitative Examples of Military Raids and PKK Attacks

	Military Raid	PKK Attack
Description	Weapons Seizure: In Lice town the Turkish military seized more than 10 tons of explosives during an anti-PKK operation. As reported on the 27th of June.	PKK killed a village chief outside the village of Bahcebasi on the night of the 26th May. The militants also burnt the car belonging to the victim.
Side A	Military Forces of Turkey (2002-2016)	PKK: Kurdistan Workers Party
Side B	PKK: Kurdistan Workers Party	Civilians (Turkey)
ACLED ID	3398	3196
Description	On August 11 one PKK militant was killed in Yumurcak village of Kiziltepe district Mardin when clashes occurred during a police operation on his home.	A car bomb exploded at around 7.30pm on May 25th at a gendarmerie miliary checkpoint in Anitli village in Midyat district of Mardin. The explosion killed one soldier and two village guards as well as the two PKK militants who detonated the bomb from inside the car.
Side A	Police Forces of Turkey (2016-)	PKK: Kurdistan Workers Party
Side B	PKK: Kurdistan Workers Party	Military Forces of Turkey (2002-2016), Gendarmerie
ACLED ID	5145	3180
	Property destruction: Between February 28 and March 2, Gendarmerie Forces destroyed 19 shelters belonging to PKK during the operations in the rural areas of Kursunlu village, Dicle district, Diyarbakir	PKK militants raided Baglica village in Artuklu district of Mardin on the 20th May and killed a village guard. Following the attack, military vehicles and ambulances rushed to the scene and eight soldiers were injured when PKK detonated explosives at the convoy. No medical personnel were reported to have been injured.
Side A	Military Forces of Turkey (2002-2016), Gendarmerie	PKK: Kurdistan Workers Party
Side B	PKK: Kurdistan Workers Party	Village Guards
ACLED ID	5515	3138

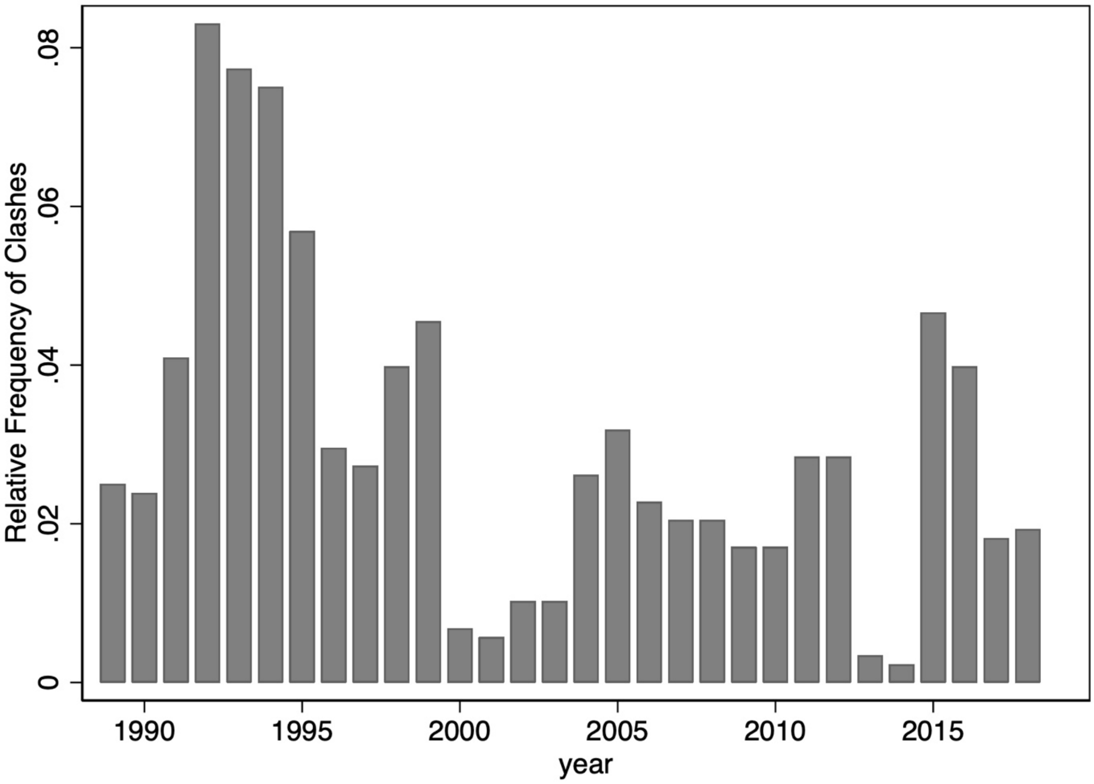
Table B2: Spatial Panel Models with Alternate Measures of Irrigation and Conflict

	(1) Clashes	(2) Clashes	(3) Recruitment	(4) Recruitment
Irrigated Area		-0.000339**		-0.000232**
		(0.000172)		(0.000116)
Years Since Irrigation	-0.00222***		-0.000644***	
	(0.000388)		(0.000236)	
Y_{t-1}	0.247***	0.253***	0.125***	0.126***
	(0.0460)	(0.0459)	(0.0374)	(0.0373)
Population	0.000210	-0.000524	0.00607	0.00595
	(0.00332)	(0.00336)	(0.00395)	(0.00396)
AKP voteshare	-0.00497	0.000336	-0.0457**	-0.0477***
	(0.0156)	(0.0167)	(0.0180)	(0.0180)
HDP voteshare	0.0760*	0.0843**	0.00225	0.00545
	(0.0393)	(0.0398)	(0.0529)	(0.0527)
Tribal	0.104***	0.0979***	0.0658**	0.0679**
	(0.0333)	(0.0332)	(0.0314)	(0.0314)
Nightlights Change	0.00100	0.00191	-0.00623	-0.00578
	(0.00601)	(0.00600)	(0.00477)	(0.00469)
Nightlights	0.00736*	0.00681*	0.00604*	0.00576*
	(0.00399)	(0.00397)	(0.00320)	(0.00314)
Roads	-0.00316	-0.00548	-0.000634	-0.000784
	(0.00316)	(0.00338)	(0.00327)	(0.00331)
SPEI	0.0191	0.0238**	0.0168	0.0172
	(0.0117)	(0.0115)	(0.0132)	(0.0131)
Slope	0.00196	0.000638	-0.000350	-0.000718
	(0.00171)	(0.00169)	(0.00207)	(0.00210)
Elevation	4.36e-05	3.99e-05	2.71e-05	2.34e-05
	(4.01e-05)	(4.04e-05)	(4.15e-05)	(4.17e-05)
Cell Area	-9.18e-05	-5.74e-05	9.21e-05	0.000118
	(9.15e-05)	(9.76e-05)	(0.000123)	(0.000124)
Ceasefire	-0.0149	-0.0120	0.0164	0.0113
	(0.0117)	(0.0120)	(0.0107)	(0.00974)
Syrian War	0.0329*	0.0198	0.00992	0.00675
	(0.0177)	(0.0177)	(0.0152)	(0.0146)
Border	-0.0156**	-0.0155**	-0.0118	-0.0112
	(0.00611)	(0.00630)	(0.00817)	(0.00820)
Observations	6,293	6,293	5,859	5,859

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

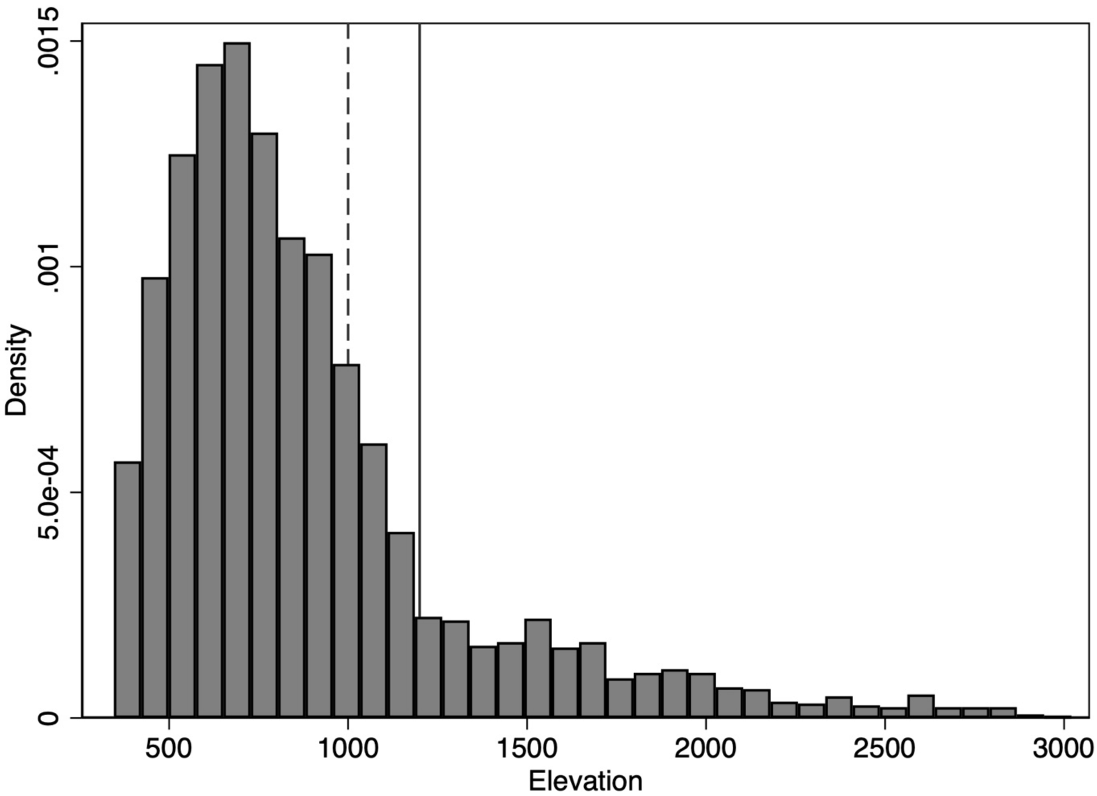
Results in this table represent modifications of the specification found in column 4 of Table 2.4. The dependent variable in columns 1 and 2 is the UCDP-derived conflict incidence measure. The dependent variable in columns 3 and 4 is an obituary-derived measure of PKK recruitment developed by Tezcür (2016). Columns 2 and 4 use irrigated area as the treatment variable, which represents the number of square kilometers of the grid cell under irrigation in a given year. Robust standard errors are reported in parentheses.

Figure B1: Frequency of Violent Clashes Over Time



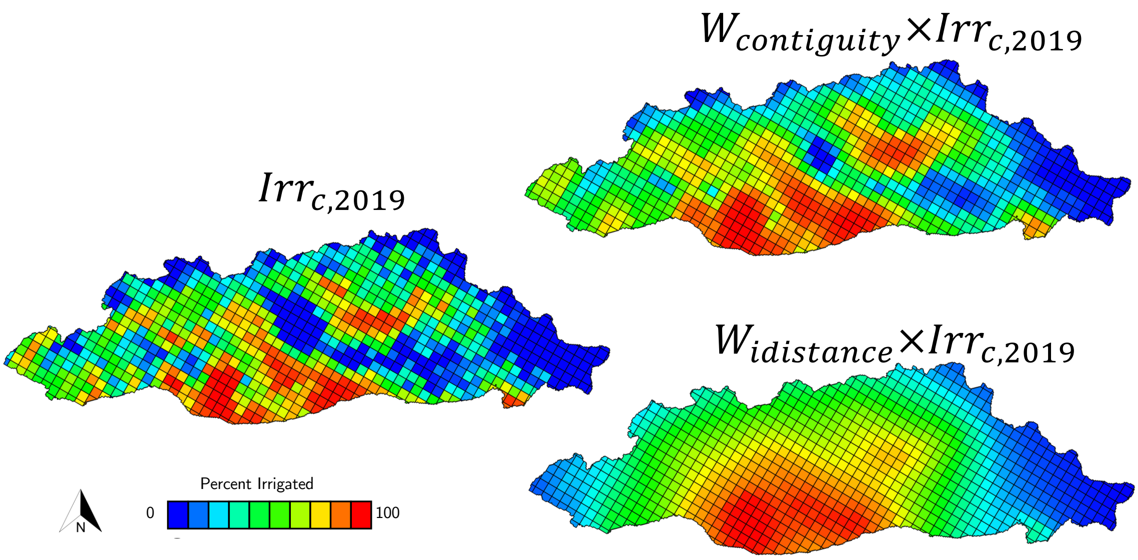
This figure shows the relative frequency of clashes between the PKK and the Turkish Government recorded in the Uppsala Conflict Data Program (UCDP) dataset. The large volume of clashes prior to the year 2000 indicates the conflict's active phase, which was followed by a ceasefire lasting until 2015.

Figure B2: Elevation Histogram of Southeastern Anatolia



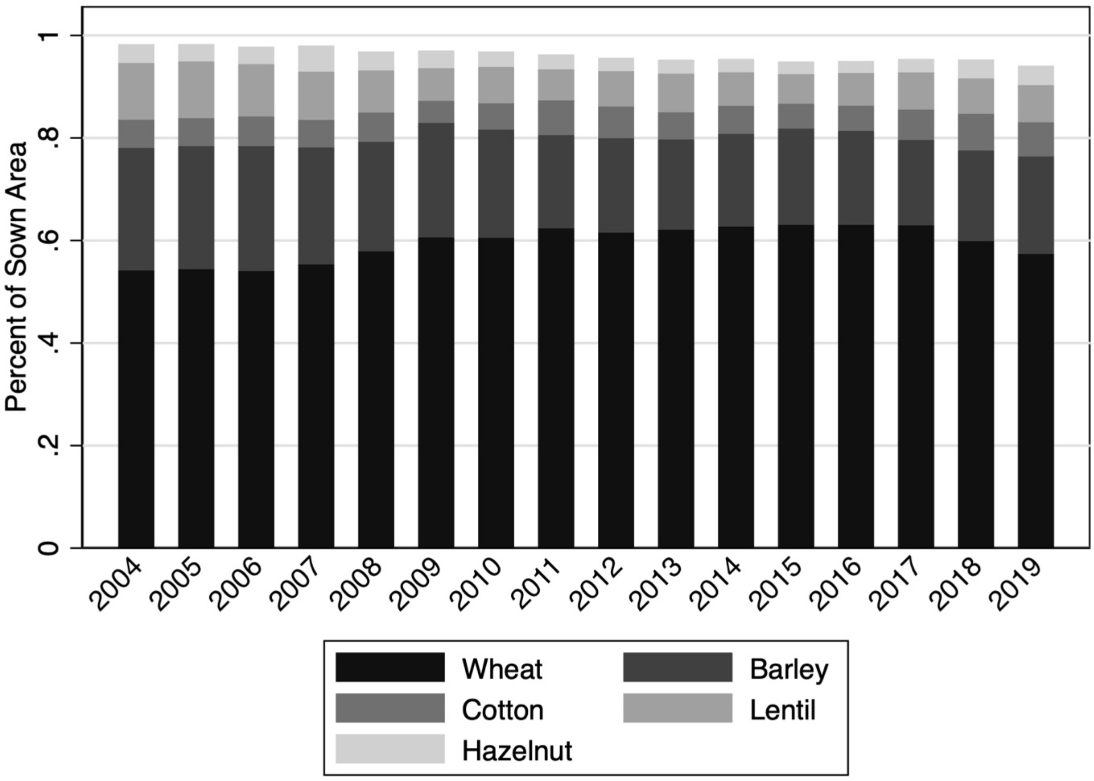
The histogram above shows the elevation profile of the study area. Mountainous areas are clearly identifiable as a long right tail in the distribution. The solid line denotes an altitude cutoff used to eliminate mountainous areas from the sample, and the dashed line shows a more aggressive geographic restriction used for robustness.

Figure B3: Effect of Spatial Weights on the Irrigation Variable



The map on the left shows the spatial distribution of the irrigation variable in 2019. The map on the top-right demonstrates the effect of an interaction between the irrigation variable and a binary contiguity spatial weighting matrix. The map on the bottom-right demonstrates the effect of using an inverse distance weighting matrix.

Figure B4: Southeastern Anatolia Crop Production Statistics



The chart above shows the relative shares of five major crops in terms of sown area. Wheat is by far the dominant crop, making up roughly 60% of sown area.

Figure B5: Central University Research Ethics Committee Approval Letter

Oxford Department of International Development
Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK
Tel: +44(0)1865 281800
qeh@qeh.ox.ac.uk www.qeh.ox.ac.uk

Head of Department: Professor Christopher Adam

5 June 2018

Ollie Ballinger
St Antony's College
Oxford

Dear Ollie

Research Ethics Approval - Ref No: CUREC 1A/ODID C1a 18-037

Project Title: The Political Economy of Agricultural Land Use Change in Southeastern Anatolia

The above application has been considered on behalf of the Oxford Department of International Development's Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethics approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

Should there be any subsequent changes to the project, which raise ethical issues not covered in the original application, you should submit details to the DREC for consideration.

Yours sincerely

Prof Laura Rival
Chair of ODID DREC

Field Research Design

Interviews were conducted with a variety of stakeholders affected by the Southeastern Anatolia Project (GAP), largely fitting into three main categories: government officials, Irrigation Union officials, and farmers. Interviews with government officials were conducted in person in Ankara, while Interviews with Irrigation Union councilors in Suruç and Batman were conducted via telephone, following a deterioration in the political situation in Turkey during fieldwork which made travel to these areas untenable. The impact of this development will be discussed further alongside the interview responses themselves. A first round of purposive sampling was carried out in interviews with farmers in order to adequately represent major stakeholder groups such as smallholder farmers, and to ensure that diverse parts of the study area are covered (Cakmak, 2000).

For case studies, two areas in particular were chosen. The first was Suruç, because a spontaneous protest began during fieldwork in response to a shortage of water; the second was Batman, because interviews with government officials suggested that the process of renationalization was being violently contested in this area. Furthermore, the quantitative analysis suggested that both of these areas were strong outliers in the overall relationship between irrigation and recruitment. Once an initial purposive sample covering these areas and representing different stakeholder groups was established, a second round of snowball sampling was used in order to provide multiple perspectives within the same communities. The overarching purpose of these interviews was to gather information about irrigation-dependent farmers' relationship with the state in the context of an insurgency, as well as the social, political, and economic factors that mediate this relationship.

Interviews were conducted between August and September 2018, during which time the Turkish government was in the process of renationalizing irrigation infrastructure. As DSİ began retaking control of irrigation schemes, many of the problems that had plagued the early DSİ-controlled irrigation projects began resurfacing: improper timing and dispensation of irrigation water led to significant crop failures, with the Suruç irrigation scheme running out of water less than midway through the summer growing season, leading to protests (Urfanatik, 2018). In Batman, a DSİ engineer was severely beaten while carrying out renationalization (Interview III). This contentious political climate allowed for an extreme test of the effect of irrigation on Kurdish farmers' perceptions of the Turkish state. Whereas previous ethnographies focused on well-performing schemes during times of political calm, the interviews conducted herein focused on particularly violent areas of Southeastern Anatolia during a time of protests and a significant intensification of the conflict with the PKK. Because the effect of irrigation on farmers' perceptions of the state is the main mechanism through which irrigation is posited to reduce PKK recruitment, conducting interviews in this political context allowed the durability of this effect to be interrogated.

The table below lists the occupation and location of interviewees, as well as the medium of the interview:

Table B3: List of Fieldwork Interviews

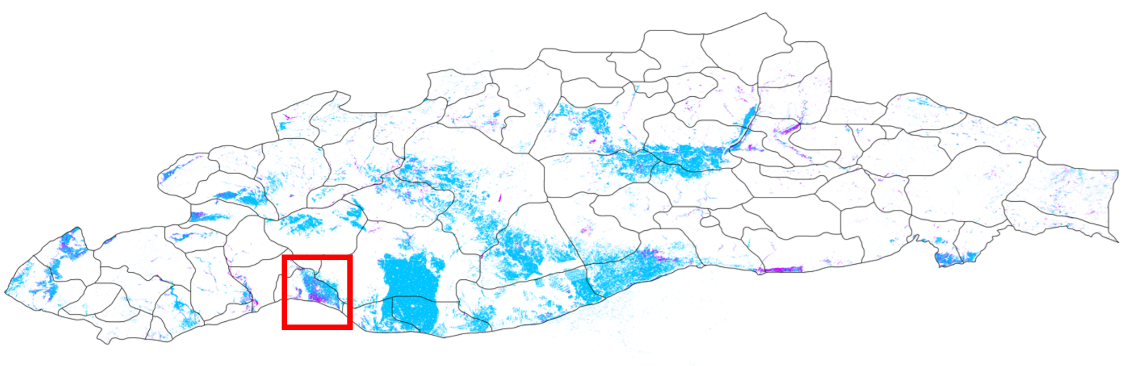
Number	Occupation	Location	Medium
I	Academic	Istanbul	In Person
II	DSİ Field Officer	Diyarbakir	Telephone
III	DSİ Field Officer	Batman-Silvan	Telephone
IV	DSİ Field Officer	Ankara	In Person
V	Retired DSİ General Director	Ankara	Telephone
VI	Turkish Court of Accounts	Ankara	In Person
VII	Judge	Ankara	In Person
VIII	SUBIRDER Official	Ankara	In Person
IX	Irrigation Union Councilor	Suruç	Telephone
X	Farmer	Suruç	Telephone
XI	Farmer	Suruç	Telephone
XII	Farmer	Suruç	Telephone
XIII	Farmer	Suruç	Telephone
XIV	Farmer	Suruç	Telephone
XV	Irrigation Union Councilor	Batman-Silvan	Telephone
XVI	Farmer	Batman-Silvan	Telephone
XVII	Farmer	Batman-Silvan	Telephone
XVIII	Farmer	Batman-Silvan	Telephone
XIX	Farmer	Batman-Silvan	Telephone
XX	Farmer	Batman-Silvan	Telephone
XXI	Farmer	Batman-Silvan	Telephone
XXII	Farmer	Diyarbakir	Telephone
XXIII	Irrigation Union Councilor	Diyarbakir	Telephone

This table reports the occupation and location of individuals interviewed during fieldwork in 2018, as well as the medium through which the interview was conducted. References to these interviews refer to the number column on the left.

Qualitative Case Study 1: Political Violence and Irrigation in Suruç

The Suruç irrigation scheme, highlighted in Figure B6, borders the Syrian city of Kobanî which has been controlled by the PKK-affiliated YPG since 2015. At the center of the scheme is the city of Suruç, which has historically sent a large number of recruits to the PKK—more than 98.7% of other grid-cells in this sample. During the summer of 2018, the mere visit of the city's incumbent Member of Parliament—an AKP deputy running against an HDP candidate widely supported in Suruç—was enough to spark multiple deadly riots (BBC, 2018). During fieldwork, this irrigation scheme ran out of water shortly after management was transferred from the local Suruç Irrigation Union to the State Hydraulics Administration (DSİ). Thus, a city with a long history of violently contesting the Turkish state's legitimacy was now surrounded by farmers whose crops lay rotting in the sun as a direct result of the state's mismanagement of water resources. Bearing in mind the quantitative association identified in Chapter 2 crop failures and violent conflict, interviewing farmers in Suruç represents a much more extreme test of the effect of irrigation on farmers' perceptions of the Turkish state.

Figure B6: Irrigated Areas in Southeastern Anatolia, Highlighting the Suruç Irrigation Scheme



The Suruç Irrigation Scheme highlighted on a map of Southeastern Analtolia.

Under the irrigation union, farmers' fields were watered every 9 days, and several villages received water at the same time (Interview VI). Under DSİ management, villages would receive water one by one, and could water their fields every 3 days

(Ibid). Because the farmers practice water-intensive flood irrigation, increasing the frequency of the rotation caused the scheme to run out of water in late July, roughly half way through the summer growing season, and right before the hottest month of the year (Interview VII). Crop failures were, in the words of an interviewee, "felâket" [catastrophic] (Interview VIII). An Irrigation Union president was quick to blame DSİ for the debacle:

> "We are farmers. We are sons of farmers. We are grandsons of farmers. We know when a crop needs water, and how much it needs. For an outsider, a bureaucrat who has never farmed, to manage our irrigation is not possible. Because of this we are suffering." (Interview IX)

Such a critique is to be expected, given the depletion of irrigation water occurred shortly after DSİ took control and altered the rotation. Yet the farmers took a very different view of the situation:

> "The president of the Irrigation Union wants the management to stay with the Union. But the farmers prefer DSİ. The work of DSİ is better. There is state protection behind DSİ, there are guarantees for timely and good quality maintenance and repair. The farmers are happy with this. We prefer DSİ." (Interview X)

Thus, while the irrigation union president emphasizes a divide between "locals" (Suruç farmers, including himself) and "outsiders" (state officials and the DSİ), the farmers frame the issue quite differently. Mehmet Faruc Cardirci, a smallholder farmer affected by the lack of irrigation water, expressed his view of the problem in this quote from a local news report:

> "There are people who do not share. There is enough water for all, but they do not respect each other. The strong does not give water to the weak. I am speaking to the President of the Republic: we cannot get a single word across here, everyone has different ideas in their heads, people are lying to each other and harming each other. Our President is the only one who will solve this trouble." (Urfanatik, 2018)

His account frames the situation among farmers as anarchic—an almost Hobbesian state of nature in which the strong may irrigate and the weak cannot. And just as Hobbes viewed the power of a Sovereign as the only solution to the war of all against all, Cardirci believes that only the President can save the farmers from each other. Despite the fact that crop failures resulted more or less directly from a state policy, the implication seems to be that the switch to 3-day rotations would not be a problem if the farmers behaved honorably. He seems to argue that by switching to a 3-day rotation, the state merely made the depletion of irrigation water possible; but it was the farmers' abuse of that system that made it a reality.

For weeks, farmers who would otherwise be tending to their crops assembled in protest outside of the regional DSİ office:

Figure B7: Photograph of Water Protest in front of Suruç DSİ Branch

Image taken from a local press report on farmer protests outside of the regional office of the State Hydraulics Administration branch in Suruç, following crop failures due to mismanagement of water resources.

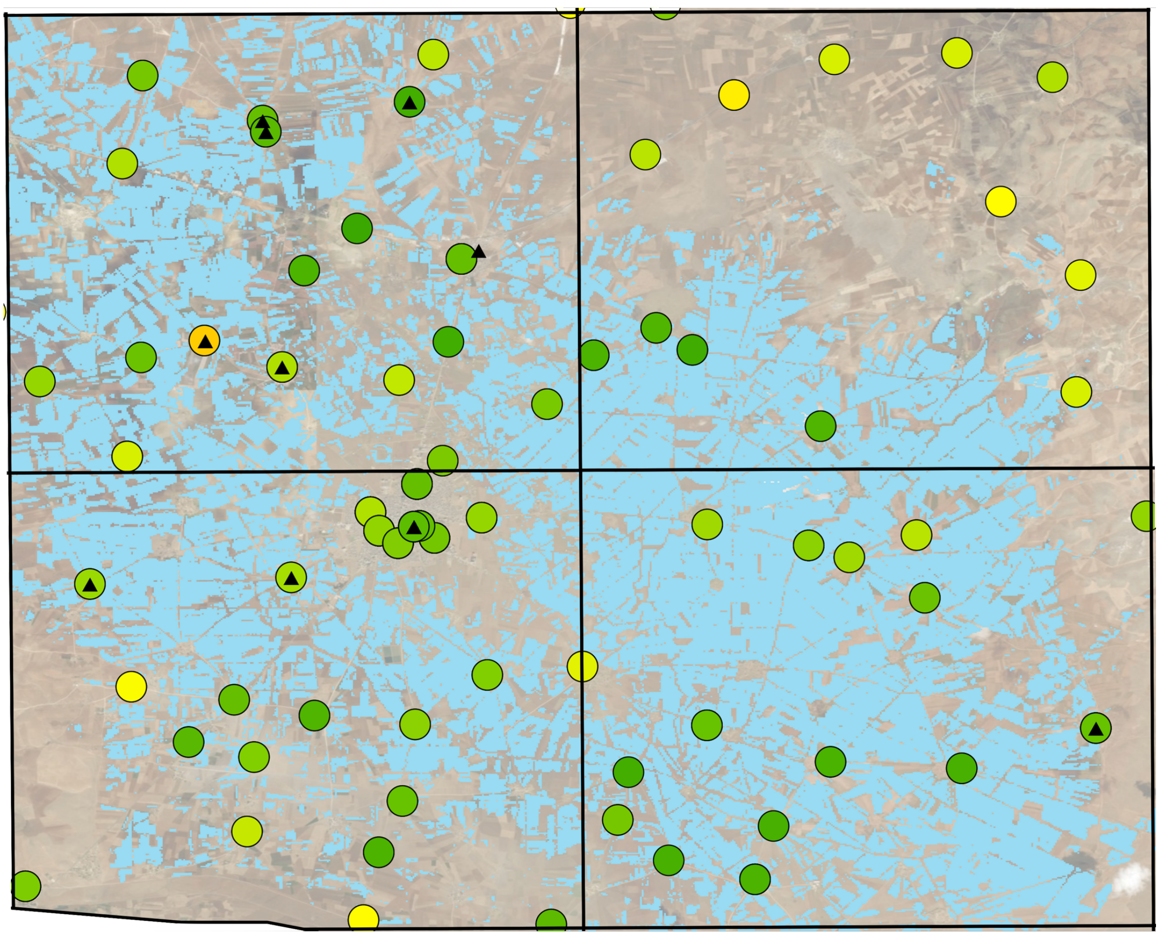
However, none of the placards being held by these farmers actually blame DSİ for their problems—quite the contrary: they are appealing to DSİ to solve them. The placards include slogans such as "The low-pressure pumps are not working", "Solve the water issue", and "We hope that DSİ will support the farmers". One placard—the one held by the man on the far left—is of particular interest. He is wearing a Cemedanî, a traditional Kurdish headscarf that was banned until the 2000s because of its association with support for the PKK (Onyebadi, 2017). His placard reads:

> "Honorable [Agriculture] Minister, Honorable Member of Parliament, Honorable Governor, Honorable District Governor, Suruç could be a monument to you" (Urfanatik, 2018).

Though "eseriniz" has no direct translation, it is most often used in the context of works of art or monuments created to memorialize gods or rulers, such as the statues of Zeus and King Antiochus on Mount Nemrut (Nemrut Eserleri). As such, this protestor directly appeals to every level of government, from the Capital to the Province to the District, promising that if they resolve the water shortage, the Suruç irrigation scheme will perform so well that it would become a monument in their honor.

This level of effusive praise and reverence towards the state is extraordinary for Suruç, which has sent more recruits to the PKK than 98.7% of the grid-cells in the sample. Only two weeks prior to this protest, the very Member of Parliament that is being addressed in the placard above—Ibrahim Halil Yildiz, a member of the AKP who represents this district—was attacked by shopkeepers during a tour of Suruç in the run up to the 2018 general election (BBC, 2018). National and government-affiliated press claimed that the shopkeepers were PKK supporters, though this is disputed in local news reports. Though Yildiz survived, the brawl left four dead (including Yildiz's brother and three assailants), eight hospitalized, and over a dozen arrested (Sabah, 2018). During the funeral procession for the assailants, police attempted to interdict the caskets on their way to the cemetery, leading to a riot in which Yidliz's rival—HDP candidate Songül Arpaçi—was shot in the head with a teargas canister (Ahval, 2018). Though Suruç later voted over 70% for the HDP, and only 28% for the AKP, Yildiz received enough support from neighbouring municipalities to win re-election. Figure B8 maps these election results on the four grid-cells that contain the Suruç irrigation scheme, with irrigated cropland highlighted in blue. Dots represent ballot boxes, with green representing HDP support, red representing AKP support, and black triangles indicating that the village has sent recruits to the PKK.

Figure B8: 2018 Election Results, PKK Recruitment, and Irrigation in Suruç



The four 10km-by-10km grid cells shown above encompass the Suruç irrigation scheme. Irrigated cropland is shown in blue, georeferenced ballot box level election results for the 2018 parliamentary election are shown as circles coloured red for high AKP vote shares and green for high HDP (Kurdish Party) vote shares. Triangles denote villages that have sent at least 1 recruit to the PKK in the time period covered by the sampple.

Though 10 villages in the scheme sent 75 recruits to the PKK, 69 of them joined before irrigation arrived in 2005. However, urban recruitment from the city of Suruç (in the bottom-left grid-cell) remained high throughout. The overwhelming support for the HDP evident in Figure B8 seems contradictory to the appeals by protesting farmers to AKP officials including the Member of Parliament and the President. However, though interviewees spoke very favorably of the state [devlet], not a single one expressed support for the government [hükümet]. This could suggest that farmers discursively separate their irrigation-related interactions with state bureaucracies such as DSİ from their party-political beliefs and preferences, allowing them to appeal to offices of the state (the President, Parliament, and Ministers) without necessarily supporting the politics of the individuals that populate them.

The juxtaposition of the assault on Yildiz and the water protest appears to be a testament to the extent to which irrigation can instill trust in the state. Suruç and its surrounding villages seem to deeply reject the legitimacy of the state, on the streets—by assaulting their visiting AKP Member of Parliament—and on the battlefield—by fostering record numbers of PKK recruits. Two weeks after the assault, the state took control of an irrigation scheme, and through mismanagement caused catastrophic crop failures. If the mere visit of a government official was enough to spark multiple deadly riots, one would expect the state's mishandling of a vital economic resource to provoke an even greater response. Yet, many of the farmers not only place the blame for the water shortage on each other, they appeal directly to the President as the "only one" who can help, and promise to become a "monument" to the state if the shortage is resolved. In one interview a farmer expressed that, though the shortage occurred under DSİ management, he would still rather have DSİ manage the water than the Irrigation Union:

> "We do not forget who dug the canals. This area was very poor. In the old days, farmers in Harran could not even produce 100 kilos of wheat. Some years there was drought, they could grow nothing and there would be famine. Now, we export wheat, we grow cotton, we have money for equipment and tractors. After the Atatürk Dam opened up, areas such as Viransehir, Harran, the Suruç plains, and the small towns and villages of Sanliurfa [province] opened up as well. So we are experiencing some problems with water shortages, but this is a new operation and a transition. This whole area will be fully irrigated again very soon." (Interview XIV)

This response frames the current shortage in irrigation water as a recent aberration in a long and constructive relationship between farmers and the state. The main temporal distinction that he draws is not between the days of DSİ management and the days of irrigation union management; it is between the pre-irrigation and post-irrigation period. Despite the fact that irrigation was introduced nearly 15 years ago, the enduring effect thereof on this individual's trust in the state appears to outweigh short term failures of the state.

The voices of farmers from Suruç—both the ones contained in my own interviews and those from local news reports—echo those of farmers interviewed in the literature, for whom "receiving renewed state attention and services has resulted in an intensified sense of belonging and loyalty as citizen subjects" (Harris, 2009: 11). Yet the present inquiry represents a far more extreme test of this loyalty. The interviews carried out by Harris (2002; 2008; 2009; 2012; 2016) were conducted in the Harran Plain, which has consistently voted for center-right parties (including the AKP) and has not sent a single recruit to the PKK. In contrast, Suruç has maintained a far more conflictual relationship with the state, and interviews were conducted in the midst of DSİ's failure in the management of irrigation water. Nevertheless, the positive effect of irrigation on farmers' perception of the state persisted.

Though this reinforces the validity of the causal mechanism linking agricultural income and conflict, it also highlights a general limitation in the scope of this analysis. While irrigation appears to have been successful in fostering better relations between farmers and the state and may even have directly deterred some farmers from joining the PKK, a substantial portion of the PKK's support base is unaffected by irrigation. Urban residents—including the shop-keepers who attacked Yildiz—continue to harbor grievances against the Turkish state, regardless of whether the farmers on the outskirts of town grow wheat or cotton.

This case study sheds greater light on the precise ways in which irrigation might deter an individual from joining the PKK when it is considered in tandem with detailed information on the recruitment process. In an interview conducted by Aytekin (2019: 72), one of the founding members of the PKK described their recruitment strategy as follows:

> "One day we went into the villages to get to know the local population and gain supporters. We knocked on a door and asked to be let in, we sat

> next to an old Kurdish man who was giving a speech. I later learned he was a 'Dede' [an Alawite preacher]. I asked him stating 'Dede, while playing your Saz [a traditional guitar style string instrument] you mentioned the region of Dersim, if you give permission I would like to mention a few things in regards to Kurdistan and Kurdish history.' The Dede then gave me permission to speak whereby I gave an effective speech stating, 'we have a land, we have a history.' [. . .] The Dede then asked who we were, we replied stating we are a new organization. That night, we then gave a speech to the men of the village between the regions of Pertek and Mazgirt. These were in the control of TIKKO [Communist Party of Turkey], but even though this was the situation, the whole village had joined our organization."

Rather than promising lavish salaries or attempting to persuade tribal leaders to pledge their allegiance, PKK recruiters seek permission from village elders to give speeches about Kurdish history, identity, and class consciousness, appealing directly to the potential recruits themselves. The success of this strategy is thus largely contingent on how receptive the audience is. The above quote suggests that it is most successful in areas that already displayed an affinity towards ideologies similar to that of the PKK. One recruit described the efficacy of this approach:

> "My main reason for joining the PKK were the PKK guerrillas. Every time I saw them, I would ask 'what are they doing?' I would get a reply stating 'they are fighting for us.'" (Aytekin, 2019: 72).

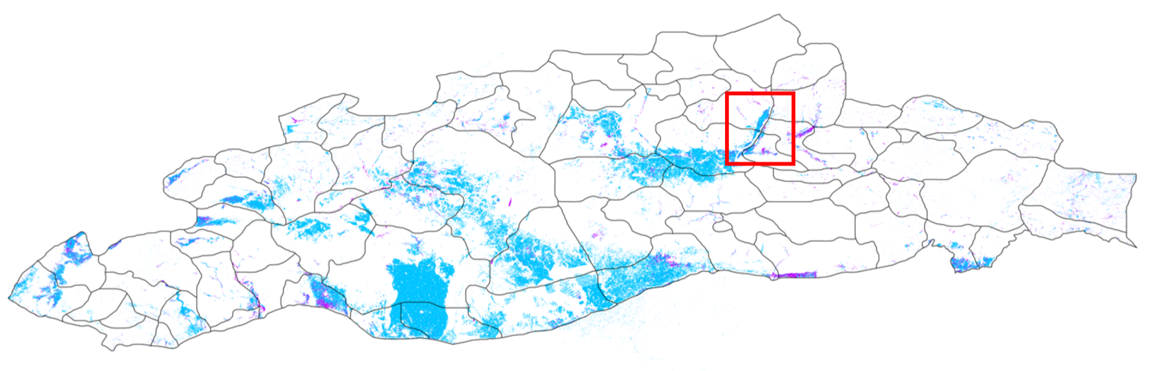
Thus, if irrigation strongly impacts how individuals perceive of the Turkish state, it likely has a strong impact on how receptive villagers would be to a guerrilla's speech. A report by Turkish Police Intelligence notes, as "the perception of the government's legitimacy declines, the insurgent movement wins over the populace's heart" (Unal, 2012: 449).

The analysis in this case study must be taken with several grains of salt. Firstly, fieldwork interviews must be interpreted in light of the circumstances under which they were produced. The interviews were carried out by phone using a snowball sampling method whereby I asked farmers to refer me to other farmers in the area. Though this process of referral established a baseline level of trust, one must still question the extent to which farmers would feel comfortable openly criticizing the government to me, a stranger on a cellphone, in a political climate in which doing so could have severe consequences. To alleviate this problem, I took care in interviewing several across several strata and stakeholder groups (government officials, Union members, farmers, etc.), and integrating local news reports to create a plurality of perspectives on the issue. The conclusions drawn in this chapter are thus tentative and would greatly benefit from more detailed follow-up fieldwork. However, qualitative evidence generally supports the causal framework and quantitative results.

Qualitative Case Study 2: Tribalism, Irrigation, and Violence in Batman-Silvan

The Batman-Silvan irrigation scheme, highlighted in Figure B9, stands out in the data as the strongest outlier in the generally negative relationship between irrigation and recruitment. Unlike Suruç, PKK recruitment remained high even after irrigation was introduced, and news reports suggest that irrigation was actually directly implicated in fueling the conflict. Data on tribal control indicated an unusually high concentration of HDP-affiliated tribes living in close proximity to an AKP-affiliated tribe. Local news reports documented a large number of deadly feuds over irrigation water, which drew involvement from the PKK and the Village Guards. In Batman-Silvan, the conflict seems to have taken on the character of local disputes—in this case, over irrigation—as hypothesized by Kalyvas (2003).

Figure B9: Irrigated Areas in Southeastern Anatolia, Highlighting the Batman-Silvan Irrigation Scheme



The Batman-Silvan Irrigation Scheme highlighted on a map of Southeastern Analtolia.

Many of the responses from interviews carried out in the area via telephone touched on themes that were suggestive of a strong tribal presence. The farmers in Batman noted that while they make more money with irrigation, their landlords are the primary beneficiaries thereof:

> "The math is like this: before irrigation, I made TL150,000 per dönüm growing wheat. Now, for one dönüm, I make TL1.5 million in income, of which TL800,000-900,000 goes to rent. A farmer who rents from a landlord

> makes about TL300,000 per dönüm. If the farmer owns the land, one dönüm of irrigated land brings TL600,000-700,000." (Interview XVI)

The exaction of excruciatingly high rent—in this case, roughly 2/3 of income—seems to fit with Guida's (2014: 179) description of tribal land tenure arrangements, whereby "Ağas became landlords, and their followers became sharecroppers".

Further interviews with government officials and irrigation union councilors indicated patterns of violence related to irrigation reminiscent of the literature on blood feuds over irrigation water. A DSİ official working on the process of renationalization in Batman-Silvan offered a succinct summary of the situation:

> "our relationship with the farmers is very bad. That area is very dangerous. Last week one of our engineers was severely beaten [darp etmiş]" (Interview III).

The president of a local Irrigation Union offered an explanation for the violence that seemed to double as a threat:

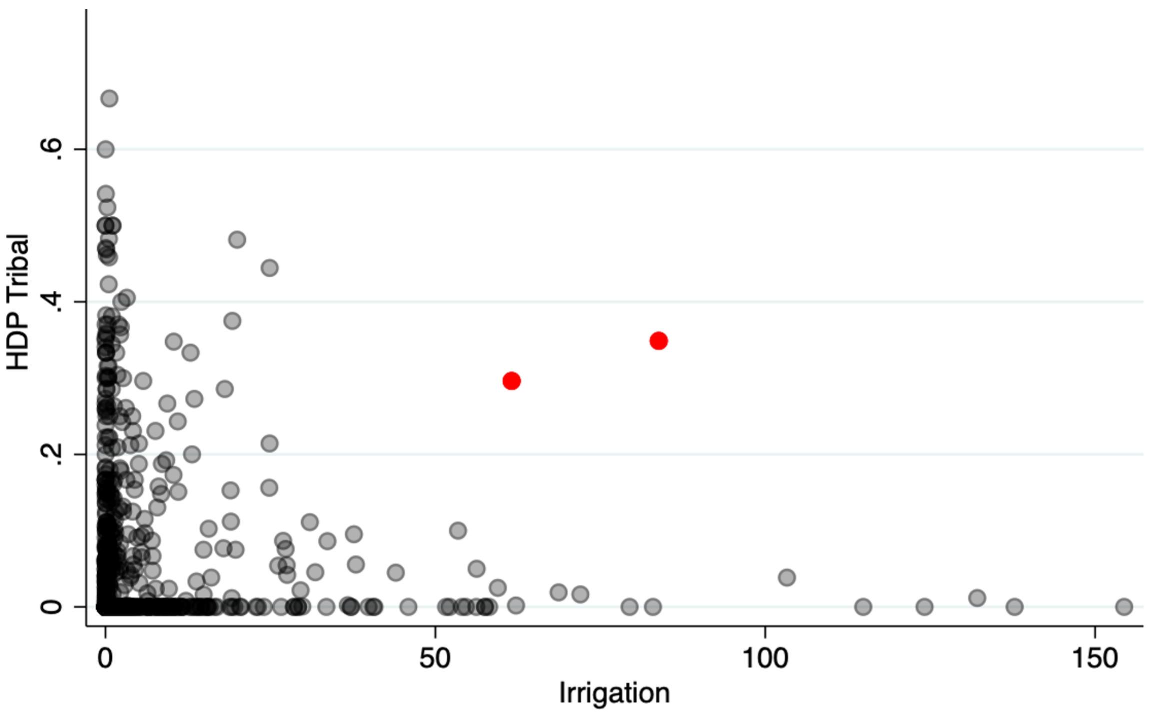
> "Allah korusun [God forbid], if one or two people start to die in each region and murders start taking place, who can do anything? There is fire all around us, there are wars, our soldiers are dying. You come here and take the bread out of the farmer's hand and give it to others? The consequences are high." (Interview XV)

The reference to "wars" and "soldiers dying" was made in the midst of Operation Olive Branch, the Turkish military incursion into Northern Syria to fight the Kurdish (and PKK affiliated) YPG. The implication seems to be that DSİ retaking control of irrigation infrastructure would open up a new front in this war, in Batman. His description of the history of irrigation management in the area bears the hallmarks of a blood feud:

> "During DSİ management, many of our villages did not receive water. I lived on the left bank of the irrigation scheme, in the central village of the plain. We were the third village on the list for water. DSİ brought

the water to our village, but when they tried to bring it to other farther villages, the farmers came out with sticks and said, 'we will not let the water pass.' When management was transferred back to DSİ [in 2018], people came and said, 'village A, this water is yours, village B, that water is yours'. But when a government worker comes, the villagers don't know these people. Farmers pump illegally and when the police comes, they say 'I'm getting water for my field, how can you stop me, my crops will die'. But then villages at the end of the canal lines started to suffer. So there were beatings and fights and people got hurt and injured. We had to get in the middle to stop the chaos." (Interview XV)

Figure B10: Relationship between Irrigation and Tribal Block Voting for HDP

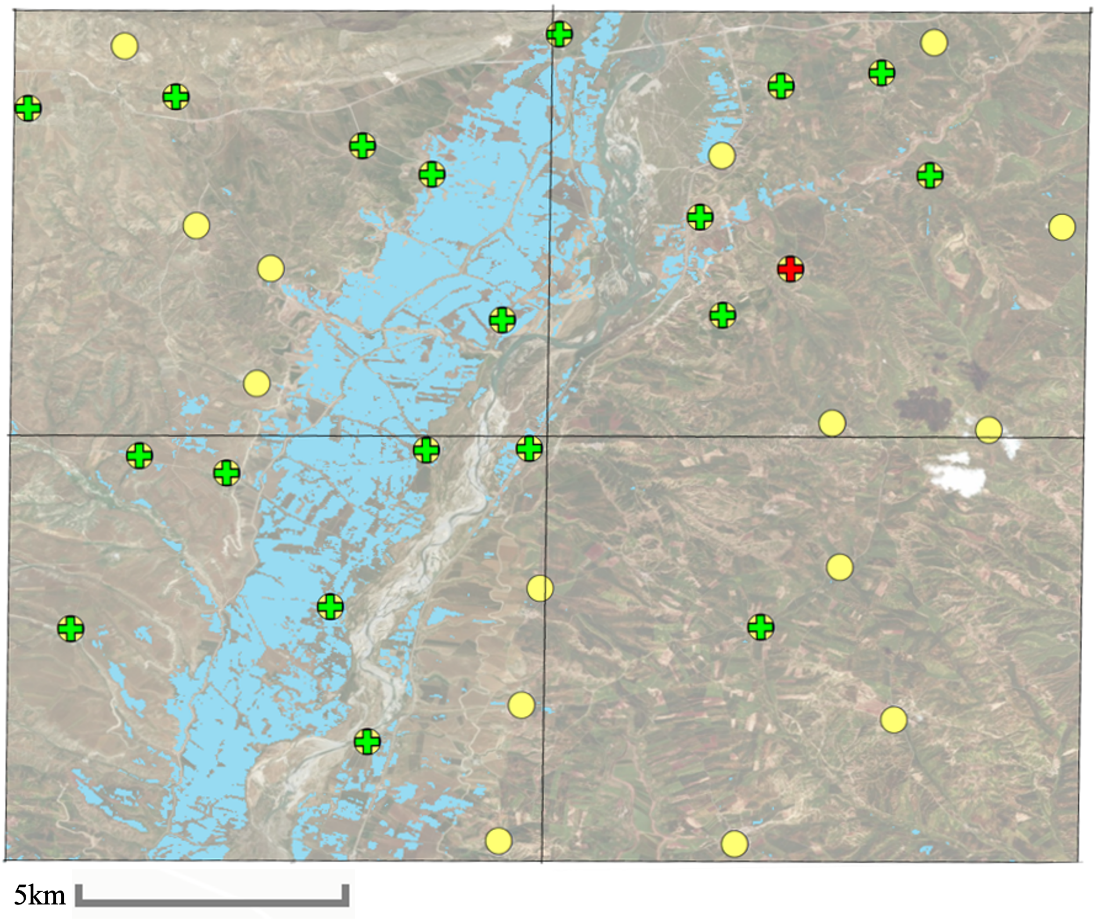


A scatterplot showing the relationship between the percent of a 10km by 10km grid cell under irrigation and the proprtion of ballot boxes displaying tribal block voting behaviour for the HDP. the two grid cells containing the majority of the Batman-Silvan irrigation scheme are clear outliers in that they have high levels of irrigation and HDP-affiliated tribal presence.

What is being described above appears to be a "tragedy of the commons" which generates violence between farmers and against state officials. Though this passage does not mention aşirets directly, it bears striking similarities to Tezçan's (1981: 64) aforementioned description of a tribal blood feud that began with the words "you are going to irrigate now? No, I am going to irrigate now!", and ended with the decimation of three generations from two clans in Kirşehir.

A re-examination of the data based on these interviews revealed that the grid-cells containing the Batman-Silvan irrigation scheme (highlighted in Figure B10) are indeed not only outliers in the relationship between irrigation and PKK recruitment—they also feature unusually high levels of tribal voting [birleşik oy]—mostly for the HDP.

Figure B11: Tribal Block Voting on the Batman-Silvan Irrigation Scheme



The four 10km-by-10km grid cells shown above encompass the Batman-Silvan irrigation scheme. Irrigated cropland is shown in blue, georeferenced ballot box level election results for the 2018 parliamentary election are shown as circles. Green crosses denote ballot boxes displaying tribal block voting behaviour for the HDP (the main Kurdish party) while red crosses indicate block voting for the AKP (ruling party).

The two grid-cells that contain the bulk of the irrigation scheme also displayed instances of birleşik oy in a relatively large proportion of the ballot boxes contained within the cells. Figure B11 superimposes the locations of ballot boxes on 2018 satellite imagery of the scheme (with irrigated fields highlighted in blue); yellow circles denote villages, and crosses denote ballot boxes in which block voting took place—green for HDP, red for AKP.

The grid-cells on the left display tribal voting behaviour in nearly every ballot box directly on the irrigation scheme. In the top right, however, birlesik oy took place in favor of the AKP, where the party won by an average margin of 96% across five elections. In the nearest village—less than a mile to the South-West—98% of the vote went to the HDP. The villages contain 198 and 207 voters, respectively. These data suggest an exceptionally large presence of HDP-affiliated tribes living in close quarters with an AKP-affiliated tribe, in the vicinity of a scarce and lucrative resource. The potential for blood feuds—defined as inter- and intra-tribal "struggle over ownership of the primary means of production" (Akpolat, 2009: 423)—thus appears to be high.

Local news reports indicate that the Batman-Silvan irrigation scheme was beset blood feuds between large clans, which induced a recruitment drive for Village Guards. In 2013, a dispute over a 500-hectare land purchase between two aşirets in Silvan led to a gun battle that left eight dead (Milliyet, 2013). The following year, the District Governor led a recruitment drive for 180 Village Guards to defend maintenance and repair work on irrigation canals and the Silvan dam, which feeds the Batman-Silvan irrigation scheme (Milliyet, 2014). Many of these Village Guards were killed, one by one, in PKK attacks over the following years (Malabadi Gazeti, 2015; Milliyet, 2016; Milliyet, 2018). Mediators who ended a similar 14-year long blood feud over an agricultural debt between the 15,000-strong Mendan and Devan aşirets in Silvan claim to have intervened in 437 feuds in the area (Hürriyet, 2007).

Kalyvas (2003: 479) suggests that "the local dynamics of conflict are more about local issues than the 'master narrative' of the war"; indeed, the conflict between the PKK and the government in the environs of the Batman-Silvan irrigation scheme seems to have taken on the character of local disputes over land, resources, and water. In this case, it seems that the positive effect that irrigation might have had on individuals' perceptions of the state was overshadowed by the increase in feuds over irrigation water, which quickly gave way to conflict between the PKK and Village Guards. Rather than writing this irrigation scheme off as a mere outlier, a qualitative analysis of the Batman-Silvan irrigation scheme demonstrates it to be a substantively important vindication of Kaylvas' (2006) call for integrating "local cleavages" into the analysis of civil war.

Georeferenced data on tribal block-voting behaviour shows significant clustering of HDP-affiliated aşirets on the irrigation scheme itself, but also the presence of an AKP-affiliated tribe only a few kilometers away. Interview responses indicated that this tribal presence had two main effects on the process of irrigation: firstly, many of the economic benefits to farmers were curtailed because of the highly unequal land tenure system. Secondly, competition over irrigation water and land between aşirets led to blood feuds that gradually became militarized, with both the PKK and Village Guards taking sides, as indicated by local news reports. Though the general trend appears to be that irrigation reduces PKK recruitment, the very opposite can be true. In Batman, the introduction of a highly lucrative common resource intersects with historical competition between aşirets over means of production, and the government's policy of deepening these divisions by selectively arming one side through the village guard system. The confluence of these factors may have resulted in the conflict between the PKK and the state taking on the character of local disputes over irrigation, thereby fueling even more recruitment.

Chapter 3

Rebellion as Complex Network: Social Ties and Hierarchy in the PKK

3.1 Introduction

The organizational structure of a militant group governs many of its fundamental characteristics including its robustness to different counterinsurgency strategies, ability to pursue negotiated settlements, and factionalization. Though the most widely cited scholarly analyses of conflict largely treat insurgent groups as unitary actors (Fearon and Laitin, 2003; Collier and Hoeffler, 1998; 2004; Cheibub et. al., 2010; Blattman and Miguel, 2010), a growing body of literature seeks to understand the complex social and political organization of these groups through the application of concepts and methods from social network analysis (SNA). The main impediment to this type of research is data availability: in a review of the literature on SNA in the study of insurgency, Zech and Gabbay (2016: 231) conclude that "The covert nature of militant groups makes them a difficult subject to study in general, a problem that becomes particularly acute when the objective is to map out the very internal structure that militants go to great lengths to conceal." In this paper, I introduce a new methodology that leverages deep learning and unsupervised clustering to automatically generate a social network graph based on co-appearance in photographs– an abundant but underutilized source of data often generated by militants themselves.

Though insurgent groups generally produce very little public data, there is one main exception; online visual propaganda is the cornerstone of many militant groups' recruitment strategies, with content ranging from combat footage, to magazines, to social media posts (Baele et. al. 2020; Dauber and Winkler, 2014). Though use of social media by rebels is likely to be uneven across the world, visual propaganda was already heavily used by insurgents over fifteen years ago. Over a five month period in 2007, U.S. forces in Iraq raided eight Al-Qaeda media labs and seized 23 terabytes of images and videos destined to be uploaded to the internet (Dauber, 2020). More recently, militants have taken to directly posting these images on social media themselves (Klausen, 2015); even loosely knit movements such as the participants in the Capitol Hill riots often document their activities on platforms such as Parler, generating over 70 terabytes of image and video data. Photographs– particularly those in which individuals appear with each other– are a rich source of relational data and are plentiful in the context of most insurgencies. Though image co-appearances have been used to construct social networks of university students and wedding-goers (Lewis, 2008; Berry, 2004), they have never been used to construct the social networks of militants. Thus, I seek to answer the following question: How can co-appearance in militant photographs be used to understand the social, political, and organizational forces that structure an insurgent group?

To demonstrate this, the following analysis uses as its only input 20,000 publicly available images published online by the Kurdistan Workers' Party (Partiya Karkerên Kurdistanê, henceforth PKK), an insurgent group primarily active in Turkey. This unstructured image data is converted into a social network graph in three steps: faces are automatically extracted from images using deep learning, nodes are generated by identifying individuals across images via unsupervised clustering, and edges are formed by linking individuals who appear together in photographs. Because "posing in a photograph with others is a deliberate act, and is generally indicative of some social connection" (Golder, 2008: 44), this process effectively measures the degree and nature of an individual's social embedding within the rebel group. The result is the PKK Image Co-Appearance Network (PICAN), a social network graph that mirrors many known characteristics of PKK. The PICAN has a number of useful analytical properties including a strong relationship between node centrality and rank, as well as a temporal dimension extracted from image metadata. These enable the identification of political factions within the PKK, a better understanding of the group's resilience to counterinsurgency tactics, and even insights on gender segregation among militants. Ultimately, this analysis demonstrates that the abundance of unstructured image data generated by insurgents can be harnessed to generate social network graphs that deepen our understanding of the organizational structure of militant groups.

This paper proceeds as follows. Section 2 reviews the literature on the application of network analytic methods to insurgent groups. Section 3 describes the image dataset derived from online PKK obituaries and discusses the performative nature of the photographs therein and their implications for the resulting co-appearance network. Section 4 outlines the three step methodological approach: extracting faces from images, clustering individuals' faces across images, and generating a network based on co-appearance. The rest of the paper analyzes the properties of the resulting network: Section 5 shows that a qualitative interpretation of different measures of centrality enables a rudimentary distinction between members of the political and military wings of the PKK, and even captures the marginalization of key figures following failed leadership struggles. Section 6 demonstrates a robust relationship between node centrality and rank within the PKK by cross-referencing nodes in the PICAN with wanted persons lists maintained by the Turkish government. Section 7 analyzes the relationship between network structure and the outcomes of various counterinsurgency tactics, with particular focus on the group's regrowth following a military defeat. Section 9 concludes the paper.

3.2 Literature Review

Though insurgent groups come in all shapes and sizes, the scholarly endeavor to understand the causes and conduct of civil conflict has largely ignored the internal characteristics of the actors involved. Sanin and Guistozzi (2010) illustrate the importance of these organizational factors using two examples: the FARC– a highly centralized "peoples' army" divided into hierarchically organized military units–and the Taliban, a loose confederation of semi-autonomous factions organized along tribal lines.

Key outcomes in both of these conflicts are directly tied to group structure, including the feasibility of negotiated settlements; successive peace deals in Afghanistan failed in no small part due to the need to satisfy a diverse array of factions within the Taliban (such as the powerful Haqqani Network), as well as the individual proclivities of their leaders (Neumann, 2014). In contrast, the 2016 peace accord with the FARC was facilitated by the fact that "leadership had to convince its membership—something much easier in a highly hierarchical organization such as the FARC-EP" (Segura and Mechoulan, 2017).

The application of concepts and methods from Social Network Analysis (SNA) to the study of insurgent movements has generated a number of theoretical propositions related to the topology of militant networks. Eilstrup-Sangiovanni and Jones (2008) argue that decentralized networks such as Al-Qaeda and the Taliban are flexible and adaptable, but face challenges with internal cohesion and decision making that can make them prone to excessive risk-taking. Stohl and Stohl (2007) contend that many ethnicity-based insurgencies such as Hamas, Hezbollah, and ETA form more centralized structures: small-world, scale-free networks which are characterized by short degrees of separation between nodes, well-connected hubs, and a degree distribution that follows a power law. These types of networks are particularly robust to the removal of nodes at random because only a small number of nodes are responsible for the overall connectivity of the network (Verma and Zhang, 2020). Due to difficulties inherent to the study of dark networks, "Empirical studies have mostly not engaged the central concerns of the theoretical literature—the relative advantages of centralized versus decentralized networks, the relationship between operational capability and network density, and the relevance of scale-free degree distributions" (Zech and Gabbay, 2016: 231). In the remainder of this section, I highlight the ways in which the current approach can help address three key challenges evident in the empirical literature: data availability, selection bias and the use of secondary sources, and ambiguous relational ties.

3.2.1 Data Availability

Though intelligence agencies have been able to apply SNA to covertly collected data such as call logs, these sources are not available to civilian researchers (Knoke, 2013). The primary impediment to the application of SNA to the study of insurgent groups and other dark networks is that these organizations are often "purposefully attempting to remain opaque" (Morris and Deckro, 2013). Yet virtually every insurgent group maintains some form of online presence, and in most contexts their use of visual propaganda has eclipsed other media in their public relations (Dauber, 2020). Images have fast become one of the main sources of data for empirical analyses of individual insurgent groups; researchers have used the vast quantities of image data generated by insurgent groups to better understand their ideological foundations (Beale et. al., 2020), shifts in their recruitment strategies (Lakomy, 2021), and processes of collective identity formation (Macdonald and Lorenzo-Dus, 2021). At the same time, a number of studies have used image co-appearance to construct social networks in other contexts, recognizing that valuable information on social relations can be gleaned therefrom. However, these have either required manual coding (Berry, 2004; Golder, 2008), or some level of structure in the image data such as the tagging of individuals in Facebook photos (Lewis et. al., 2008). In the context of militant photographs, image data is almost certain to be unstructured and the manual recognition and tagging of thousands of unknown individuals is unfeasible.

By leveraging facial recognition technology and unsupervised classification to automate the generation of an image co-appearance network, the approach taken herein is far more scalable than manual coding, and can utilize completely unstructured image data. Both of these features are critical given that image data in the context of insurgent groups are likely to be both voluminous and unstructured, with basic information including the number of individuals contained in the photographs likely to be unknown. Furthermore, the image processing pipeline allows for the collection of a rich array of node-level attributes through the extraction of image metadata including location, date, camera type/serial number, as well as apparent age and gender estimation using deep learning. The code used to turn unstructured image data into a co-appearance graph has been published as an open source Python package.

Using three functions, researchers can easily extract faces from images, cluster faces based on similarity, and generate a graph based on co-appearance. The package features a flexible implementation that can be optimized according to the nature of the image data at hand; this includes options to extract metadata and perform apparent age and gender estimation, enabling attribute-based analysis (e.g. homophily, ERGMs, clustering) and where datetime metadata are available, analysis of the network's evolution over time.

3.2.2 Selection Bias and Secondary Sources

Due to the extreme scarcity of primary source data generated by covert organizations, the vast majority of existing SNA studies of insurgent groups rely on secondary data, including newspaper articles (Rodriguez, 2005; Krebs, 2002; Metternich et. al., 2013), legal documents (Jordan, Mañas, and Horsburgh, 2008), and public statements (Magouirk, Atran, and Sageman, 2008). This is not only time consuming– thereby limiting scalability and resulting in smaller sample sizes–it introduces a measure of subjectivity into the analysis. Koschade's (2006: 567) social network analysis of the Jemaah Islamiyah cell responsible for the Bali bombings in 2002 contains only 17 nodes, connected on the vague basis of having had "numerous weak contacts over the period in question" or having resided together. This requires significant prior knowledge of the actors involved, limiting the novelty and utility of insights gleaned from the analysis. It also imposes an extreme degree of selection bias that effectively precludes the inclusion of anyone below the highest echelon.

Selection bias is generally less severe in the rare studies that rely on primary data. Using a trove of thousands of ISIS job applications, Edgerton (2022) is able to study suicide bomber mobilization and kin/peer ties at the level of individual recruits. Indeed, using ego networks constructed through interviews, Stys et. al. (2021) found that even complete data on "covert" group membership neglects important actors, as the liminal space between the rebel group and society is populated with individuals (e.g. demobilized combatants) who can act as brokers or fulfill other important roles. The use of photographs does not fully solve the issue of selection bias: some individuals may be systemically important but camera shy, while others may simply be photogenic.

The extent of this bias depends on the context in which pictures are taken, and is discussed further in the following section. However, social networks created using image co-appearances are likely to yield a far more complete picture of a militant group than those created with secondary sources; lower level commanders and foot soldiers are unlikely to appear in news articles and legal documents, but may well appear in a group photograph, footage from a patrol, or pictures of a ceremony. Thus, a key advantage of the present approach is that it utilizes primary data in the form of images published by militant groups, and can yield a more complete picture of an organization.

3.2.3 Ambiguous Relational Ties

A significant methodological concern in the application of network analytic methods to the study of militant groups is "the use of implicit, poorly characterized relational ties in many studies. Some analyses use relational data coded to cover a wide range of social exchange" (Zech and Gabbay, 2016: 231). In a study of the Provisional Irish Republican Army (PIRA), Gill et. al. (2014) create a network of relational ties between individuals including marriage, friendship, blood relation, and co-involvement in PIRA activities. Even within relational ties of the same type (friendship, kinship, cohabitation), there can be substantial variation in the weight of the tie that is rarely accounted for (Pedahzur and Perliger, 2006). If the ties between nodes encompass a diverse range of social interactions, network statistics and hypothesis tests derived from an aggregate quantitative analysis of these edges lose conceptual clarity.

Co-appearances in images have been used to construct social networks in several influential studies, and are generally recognized as signalling significant social ties between individuals. Berry (2004) argues that the use of co-appearances in wedding photographs provides a less subjective measure of interracial friendships than traditional measures thereof including surveys. In the context of facebook photos, Lewis et. al. (2008) suggest that co-appearance reflects a relatively high level of positive affect between the individuals involved, as well as a desire to be socially recognized with the individual. Though the process underlying when and where militants take pictures is governed by specific social practices (e.g. training academy graduations, platoon photographs, celebrations), co-appearance acts as a reliable proxy for the existence of personal contact between individuals. This, in turn, plausibly measures the extent of an individual's social embedding within an organization: if the same person is persistently present in group photographs in a militant organization, there is likely to be a reason.

Hogan's (2010) "Exhibitional Approach"– and specifically his elaboration of boyd's (2007) concept of "collapsed contexts"– provides a useful framework through which to understand the nature of co-appearances in a performative network. In the same way that friendship on a social networking site can denote anything from neighbours to family members, co-appearance in images encompasses a number of possible social relations between individuals: commanders and subordinates, comrades, friends, and acquaintances may all appear together in photos. As such, the social practices that govern when and where pictures are taken are the primary force that shape the nature of a co-appearance network. For example, images collected by a wedding photographer and a CCTV camera would likely generate two very different co-appearance networks of the same wedding. The former would deliberately overrepresent the bride and groom and under-represent the catering staff; the latter would connect people who spend more time talking to each other rather than who they choose to pose with. Thus, while a key strength of the automatic generation of a co-appearance network is that it can be applied to a wide range of unstructured image datasets, an integral part of the analysis of the resulting network is a qualitative understanding of when, where, and for whom pictures are taken.

3.3 Data: Militant Photographs

When a member of the PKK is killed, an entry for the individual is added to the group's obituary website. Each obituary features photographs of the recruit, including a high definition portrait. These photographs appear to have been taken for a wide variety of reasons. The most common type of photograph is a posed portrait of the fallen militant. Others feature militants posing next to a commander. Many appear to be taken at various ceremonies and celebrations including funerals, training academy graduations, or during Nowruz. Photographs are not taken during executions of suspected traitors, nor are those who suffer this fate given the honor of an obituary. As such, this image dataset represents a collection of the moments that PKK members and photographers have decided to capture, and the social interactions between the individuals therein. This section explores different aspects of this performance and how they are likely to manifest themselves in a co-appearance network.

3.3.1 Posed Photographs

19,115 images were downloaded from 2484 individual obituaries dated between 2000 and 2021. An interactive visualization generated using the entire image dataset used herein allows readers to qualitatively explore the photographs. The selection of many of these moments and individuals are doubtlessly curated to present an “idealized front” which furthers particular narratives about the PKK. For example, the representation of women in these images is not only a manifestation of the PKK’s feminist ideological tenets, but is an integral part of the group’s public relations and recruitment strategies. Wood (2019) notes that “The image of a smiling young woman holding an assault rifle has become a common feature of media reports about the Partiya Karkerên Kurdistanê (PKK)”. Indeed, women are slightly over-represented in these images: on average, women appear in 6.1 images, while men appear in 5.2. A T-Test suggests that this gendered difference in mean photograph appearances is significant (t=5.18; p¡0.001). Because women appear in more photographs than men, they are likely to appear more central in the co-appearance network.

Similarly, images of insurgent leaders are inherently performative. Though a rebel group may want to avoid exposing their command structure in photographs, images of the leaders amongst the rank and file rather than cowering in caves serve to cast them as both fearless yet humble. If this were the case, we could reasonably expect a relationship between an individual’s seniority and the number of co-appearances: a foot soldier may only appear with one or two commanders, but a commander would appear with many foot soldiers. However, the ways in which hierarchy is discernible from these images may not be consistent at all levels of the command structure; despite being a pivotal figure in the Cuban Revolution, Juan Almeida Bosque’s likeness is nowhere near as ubiquitous as that of Che Guevara. Individuals who are systemically important but not particularly salient could be under-emphasized in an image co-appearance network.

Figure 3.1: A Photoshoot with the PKK's Leader

These images were taken between 07:24 and 12:41 on 05/05/2014, using a Canon EOS 70D with the serial number "43021010637". They depict Murat Karayilan, the current leader of the PKK, indicated by the red box. They also feature Ekrem Güney, a mid-level commander, indicated in yellow.

The six images in Figure 3.1 illustrate how the portrayal of leaders as fearless and humble generates a positive relationship between image co-appearance and rank. These images were taken between 07:24 and 12:41 PM on May 5th, 2014, using a high-end camera (Canon EOS 70D) with the serial number "43021010637". They were the only images taken with this camera.

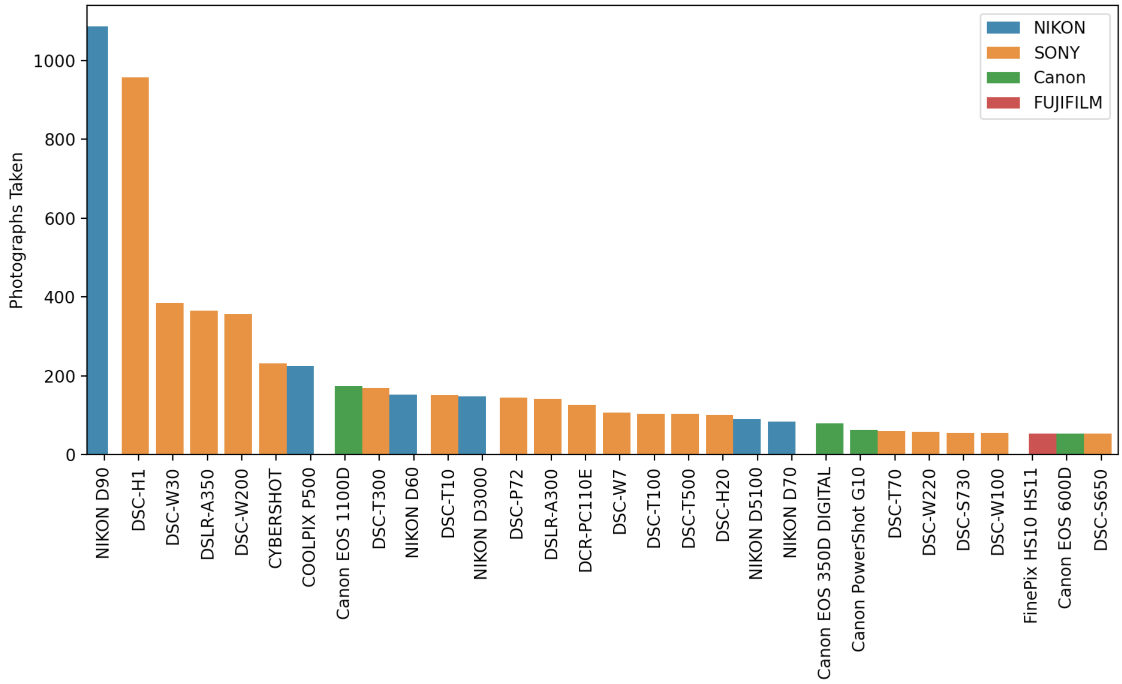
An older man is present in all six pictures. In the top three photos he poses with the same man (who appears to be in his 50s), including one in which they are seated with a third man, sharing tea and talking. In the three remaining photos, he poses next to a different young man each time. The older man present in all six images is Murat Karayilan, the 67-year-old current leader of the PKK. The middle-aged man is Ekrem Güney, a mid-level commander who was killed in 2016. The three young men appear to be foot soldiers. A co-appearance network generated from these images would see Karayilan with seven edges, Güney with four, and the foot soldiers with one edge each. This type of posing links rank with co-appearance in such a way that it not only separates the leadership from the infantry, but correctly orders the leadership in terms of rank. Though these images appear to be from a photoshoot– all taken with the same camera within hours of each other–this general principle reasonably extends to other social practices captured in these photographs in which posing is common, including ceremonies and holidays.

3.3.2 Candid Cameras

Though many of these images are staged and curated to advance certain narratives of the PKK, the majority appear to be relatively candid. There are two likely reasons for this. The first is that paradoxically, candid pictures in this context make for better propaganda. In a study of online representations of martyrdom during the Arab Spring, Halverson et. al. (2013) found that the depiction of martyrs as being relatable was crucial in generating the "imagined solidarities" that translate the collective identification with the martyr into support for a social movement. As an online repository of PKK obituaries, the emotive force that these images draw on is empathy– they depict seemingly ordinary people eating meals together, playing backgammon, laughing; ordinary people who were killed by the Turkish state. Because empathy relies on one's ability to place oneself in an other's position, candid images in this context are far more affective than staged propaganda photos.

The second reason many of these images appear candid is that the PKK publishes more traditional and overt propaganda elsewhere: not only is this obituary website not the primary way in which the PKK disseminates propaganda to the wider world, it's not even the PKK's main website (hpgonline.com). The group also publishes several magazines (Serxwebun and Berxwedan), maintains social media accounts on various platforms, and even broadcasted an international satellite television channel (Roj TV) which has been replaced by an online streaming platform (Gerîla TV). The PKK disseminates a wide range of content on these platforms, ranging from combat footage to poetry. Thus, while the PKK makes extensive use of online visual propaganda to portray itself to a public audience, this is not the primary function of the obituary website from which the images used in this study are derived.

Figure 3.2: Cameras Used by the PKK



The Y axis shows the number of pictures taken with each type of camera. The plurality are taken with a high-end cameras, but the majority are taken with cheaper models.

Metadata extracted from these photographs contains the make and model of the cameras employed by the PKK, supporting the proposition that many of these images were taken by amateurs. Figure 3.2 shows the number of photographs in the sample taken with different types of cameras. Though the plurality of these images were taken using professional DSLR cameras such as the Nikon D90 or the Sony DSC-H1, the majority were taken with cheaper digital cameras accessible to amateurs such as the Sony DSC-W30 and W2000. The notion that many of the images taken with these cameras are "candid" is not meant to imply objectivity– the deliberate presentation of candor is itself a form of performance. Rather, it simply emphasizes different actors and social relations: this style of photography likely increases the embeddedness of low-ranking individuals in the network by including group pictures of friends and comrades.

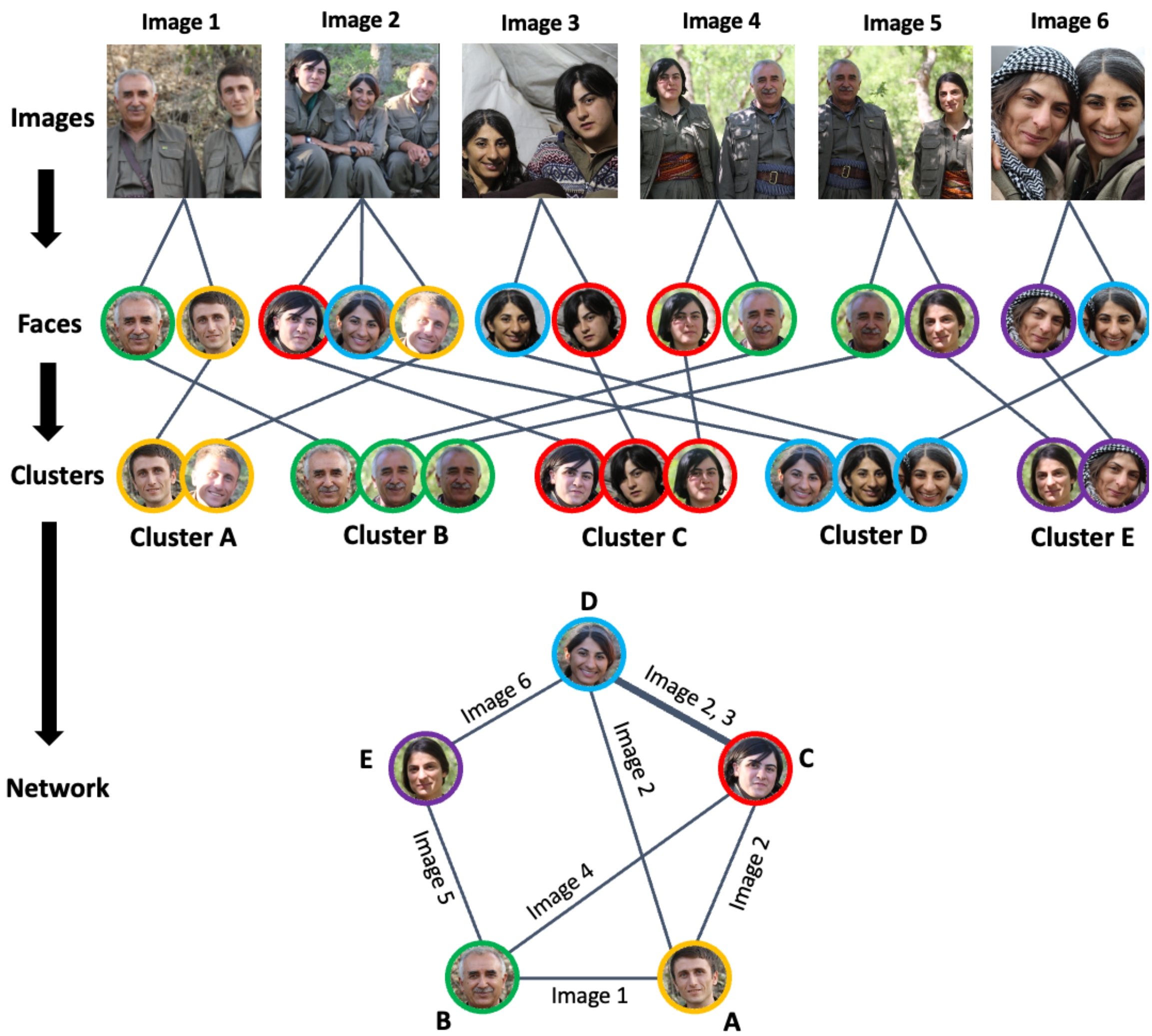
3.3.3 Privacy

Finally, the collection and analysis of image data in this way may raise valid concerns over the right to privacy in online research. In a critique of Lewis et. al. (2008), Zimmer (2010) articulates a set of ethical concerns related to the use of Facebook to create detailed profiles of college students that were anonymized but subsequently re-identified through triangulation. In particular, Zimmer highlights the use of in-network Research Assistants collecting personal information that was visible to Facebook friends but not to the general public, without their consent. The present analysis differs in two substantive ways. The first main distinction involves the feasibility of acquiring consent in the first place, as the vast majority of individuals who are included in this analysis are deceased. Still, the question remains whether living insurgents have a reasonable expectation of privacy in propaganda photos posted publicly. In this sense, the current research design is much more akin to that used by Berry (2006), who sampled publicly available wedding photos without informed consent from those who appear therein. The nature of militant photographs could be argued to entail an even lower reasonable expectation of privacy than the use of wedding photos; whereas weddings are by nature intimate and private, the very purpose of online propaganda photos is for the individuals therein to be widely seen. Indeed, of the relatively few individuals in the present analysis who are still alive, several are virtually public figures– see Murat Karayilan's headshot in the New York Times. Nevertheless, researchers applying this technique must be mindful that there may be a reasonable expectation of privacy in photographs acquired from other sources, such as on social networking sites.

3.4 Methodology

Following collection, the unstructured image data is converted into a social network based on co-appearance in three main steps. First, I use a Residual Neural Network to extract facial encodings from images. I then use a graph clustering algorithm to identify individuals across images, generating a unique cluster of faces for each individual. Finally, I construct a weighted graph in which each cluster is a node, and each co-appearance constitutes an edge. These three steps are illustrated in Figure 3.3.

Figure 3.3: Automatic Image Co-Appearance Network Creation



There are three steps in the automatic generation of an image co-appearance network; extracting faces from images, identifying the same face across images, and linking faces based on co-appearance

3.4.1 Extracting Faces

The first step in this analysis involves extracting faces from images. The approach taken herein relies on the FaceNet system developed by Schroff, Kalenichenko, and Philbin (2015), which converts face images to a "compact Euclidean space where distances directly correspond to a measure of face similarity". These 128-dimensional face embeddings can be used to store, compare, and cluster faces based on similarity. A pre-trained Residual Neural Network with 29 convolutional layers based on He et.

al. (2015) was used to detect faces, generate embeddings, and perform apparent age and gender estimation. The model was accessed via the Python distribution of Dlib, which reaches an accuracy of 99.38% on pairwise facial recognition using the standard "Labeled Faces in the Wild" (LFW) benchmark dataset (King, 2009).

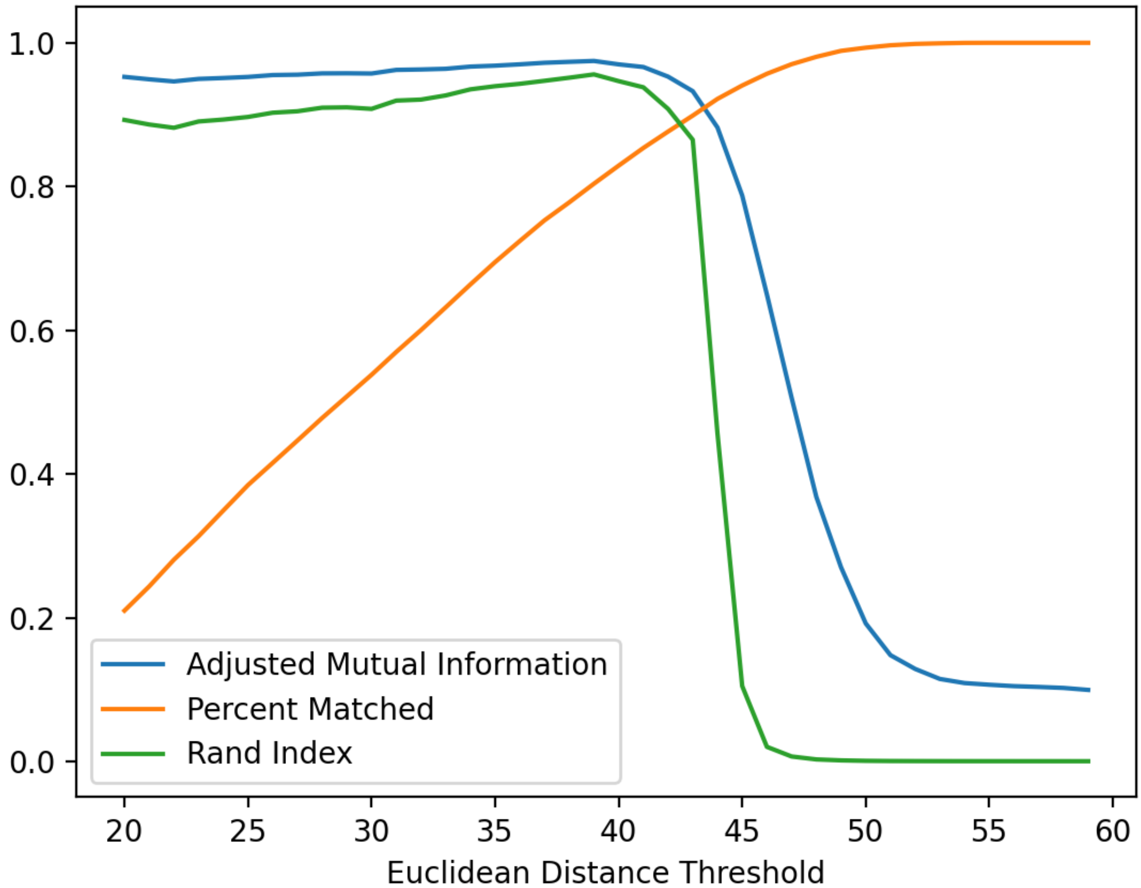
In total, 23,496 faces were extracted from 19,115 images. Cropped face image tiles are generated to enable manual verification, and facial embeddings are stored for clustering. Inspection of face tiles reveals a number of interesting artifacts, including the extraction of artistic representations of Abdullah Öcalan in the form of flags, paintings, and even lapel pins; Samples of these clusters are provided in the appendix. These are discarded.

3.4.2 Clustering Faces

Once faces are extracted, they are clustered based on similarity to identify individuals across images. The accuracy of the clustering process is assessed through the construction of a labeled subset of the data. Though images from a given obituary will primarily contain the face of the deceased individual, they also frequently contain group photos in which other faces are present. However, if an image from an obituary only contains one face, this is almost always a portrait of the obituary's subject. Thus, by restricting the sample of images to only those that contain one face, a labeled subset of the data is created.

The choice of the clustering algorithm is informed by a number of constraints in the data. Because some of the people who appear in the images are still alive, the number of unique individuals (and therefore, clusters) is unknown, precluding the use of algorithms such as k-means or spectral clustering. A number of unsupervised clustering approaches have been applied to the task of face clustering in recent years, including among others Agglomerative Hierarchical Clustering, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), and Graph-Based Clustering. Graph-based clustering has recently been shown to achieve high accuracy in face clustering tasks (Chang et. al. 2019). This approach begins by creating an undirected weighted graph $G = (V, E)$ in which each vertex (V) is a facial embedding and each edge (E) is weighted according to the euclidean distance between vertices. In the construction of this input graph, only edges whose weights are below a certain euclidean distance cutoff are included. The optimal cutoff parameter is selected based on the results of tuning shown in Figure 3.4.

Figure 3.4: Unsupervised Clustering Accuracy



Effect on accuracy of tuning the euclidean distance parameter in graph clustering. The optimal threshold is 0.39

Once the input graph is created, the Chinese Whispers (CW) algorithm is used to partition the nodes into clusters. CW functions by first assigning a unique class to each node. Through an iterative process, each node inherits the strongest class in its local neighborhood, which is the class with the highest sum of edge weights to the current node (Biemann, 2006). The CW algorithm achieves a remarkable degree of accuracy in the current dataset.

Maximum accuracy is achieved when the euclidean distance threshold is set to 0.39. If the cutoff is too low, false negatives increase as faces are not sorted into clusters. If the cutoff is too high, the number of false positives increases. The Rand

Index, which measures the similarity between a set of labeled and predicted clusters, achieves a value 95.6% under optimal conditions. The Adjusted Mutual Information index provides a similar assessment of clustering performance, but is more robust to unbalanced data with both very large and very small clusters; it achieves a value of 97.4% under optimal conditions, with 80.3% of face tiles sorted into clusters. The relatively high proportion of faces that are not sorted into clusters largely reflects poor image quality under certain conditions; large group photographs where some individuals are in the distance, low resolution images, or those taken with old cameras present a challenge to facial recognition.

The quantitative accuracy statistics provided above are enabled by the fact that images are sorted by obituary. As a final step, a mosaic image composed of all of the face tiles in a given cluster is created, enabling a visual assessment of clustering performance where image data are completely unstructured. Figure 3.5 shows an example of the face tile mosaic generated for Bahoz Erdal, the leader of the PKK's armed wing. The clustering process manages to identify him across an impressive range of lighting conditions, pose variations, image resolutions, ages, and hair styles without any false positives.

Despite the high accuracy achieved in the current dataset, Buolamwini and Gebru (2018) note significant disparities in the accuracy of facial recognition algorithms when applied to phenotypic subgroups and women, with error rates of up to 34% for darker skinned females compared to 0.8% for lighter males. Kurdish people are Caucasian and the PKK is highly ethnically homogenous, enabling high levels of accuracy. When I applied the present methodology to the Labeled Faces in the Wild dataset– a widely used benchmark mentioned explicitly by Buolamwini and Gebru– I found significant racial disparities in accuracy. The cluster generated for George Bush was extremely accurate, identifying him in 542 images– exceeding the number of labeled faces of George Bush (530) in the dataset itself– by picking up his face in the background of 12 additional images. There were only six false positives, corresponding to accuracy of 98.9%. In contrast, the third largest cluster contained 253 images of 83 different Asian people of both genders. Accuracy was significantly improved within phenotypic groups by varying the distance threshold, but the methodology presented herein would

Figure 3.5: Auto-Generated Cluster for HPG Commander Bahoz Erdal

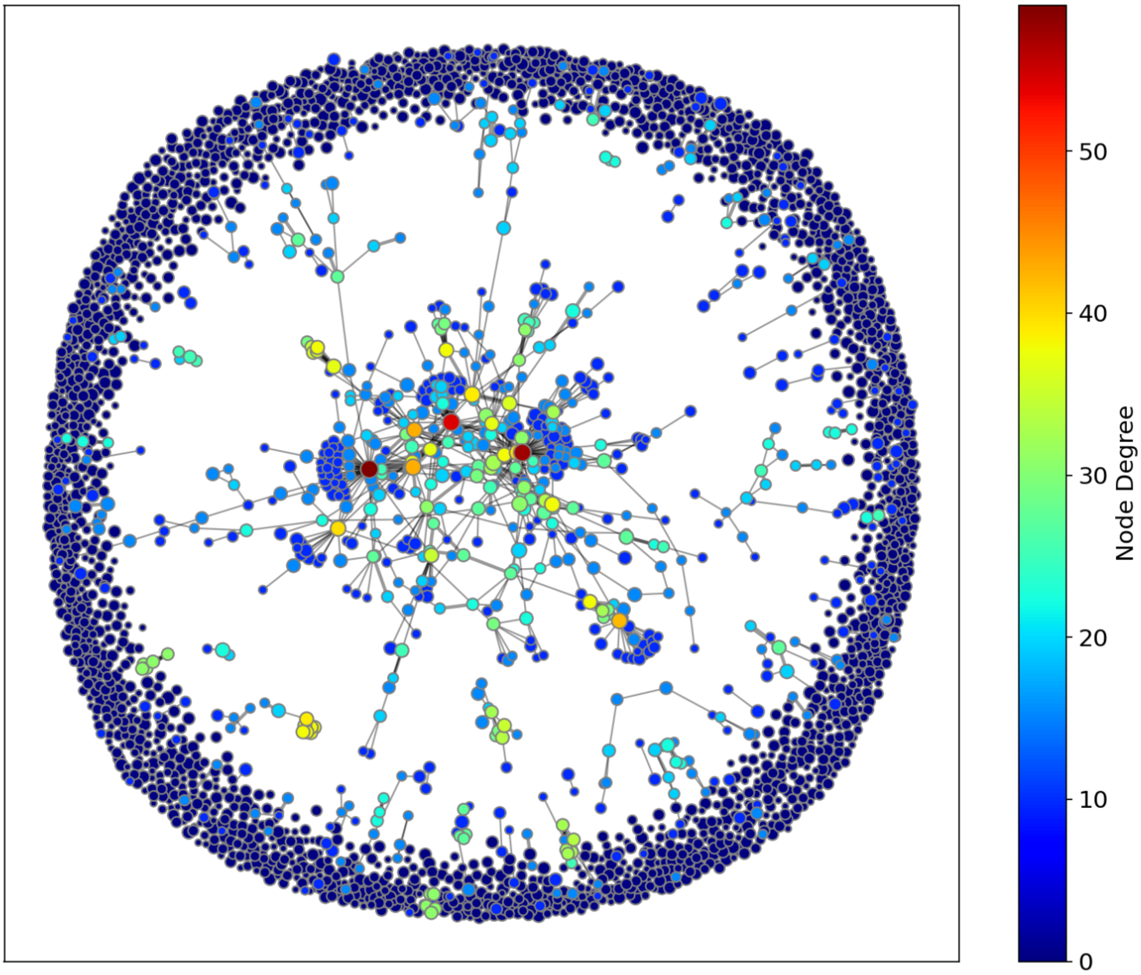
The clustering algorithm identifies Erdal in 72 unconstrained images. He is identified across a number of different lighting conditions and age ranges.

perform significantly worse on groups that are highly ethnically diverse or that consist mostly of darker-skinned individuals.

3.4.3 Generating a Network

Having identified individuals across multiple images, a network $G = (N, E)$ is generated wherein each node N is an individual, and each edge E connecting nodes indicates that the individuals appeared in a photograph together. The PKK Image Co-Appearance Network (PICAN) is weighted, meaning that an edge linking two vertices will take on the value of the number of co-appearances between the two individuals. For example,

Figure 3.6: PKK Image Co-Appearance Network (PICAN)



This figure displays a co-appearance network generated from nearly 20,000 images published online by the PKK. Each node is a cluster of faces corresponding to an individual (see Figure 3.4). Edges link nodes if the individuals appear together in a photograph, and are weighted by the number of co-appearances. Nodes are sized based on the number of pictures, and colored according to the number of co-appearances.

if individuals a and b appear together in one photograph, edge $(a, b) = 1$. If individuals c and d appear together in 10 photographs, edge $(c, d) = 10$. The degree k_i of node i reflects the number of edges linking that node to others. These edges are undirected.

In Figure 3.6, a node's size is dictated by the number of images of an individual, while the node's color reflects the number of co-appearances. The network has 2999 nodes and 991 edges. A large number of these nodes are isolates, meaning that they have no edges (k=0), making up the ring of blue nodes around the graph. These isolates derive from obituaries in which there are either a small number of images or which only contain portraits. The largest connected component (LCC) comprises

491 nodes, and is represented as the densely connected object in the center of the graph. Though there are also a number of smaller connected components, subsequent analysis will focus primarily on the LCC. Having generated the co-appearance network, the rest of this paper analyzes the properties of the PICAN in reference to external information about how the PKK is organized.

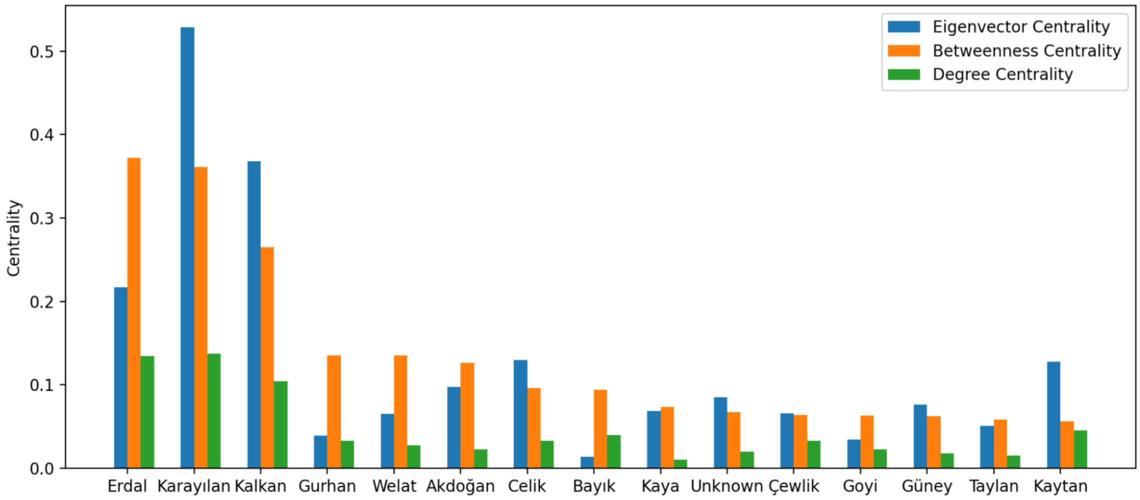
3.5 Functional and Factional Divisions

Golder (2008: 44) argues that an image co-appearance network "is a representation of a social network, not the network itself; the network itself manifests itself through the relationships of its constituent members, and is ever-changing and largely indeterminate. Indeed, a representation of a social network is valuable only to the extent that it closely reflects the composition of the actual network." The "actual network" in this case is the hierarchical organizational structure of the PKK, while the PICAN is merely a representation of the social interactions of its members. Nevertheless, the performative aspect of picture-taking in insurgent groups– particularly the phenomenon of posing with higher ranking individuals– may allow the latter to act as a proxy for the former.

There are multiple mechanisms through which node degree in an image co-appearance network could act as a proxy for seniority. The most parsimonious explanation is that most high-ranking members of the PKK have simply been active in the field for longer. The longer an individual is active, the more opportunities they have to appear in images with others. Another likely reason linking node degree and rank relates to the occasions in which pictures are taken. As previously discussed, many photographs are taken at ceremonies in which high-ranking members are present such as training academy graduations. New recruits often pose with commanders during these ceremonies, thereby generating a large number of edges in the PICAN.

The performative nature of the photographs enables a rich qualitative interpretation of network statistics. Degree centrality simply measures the number of neighbors that a node has, which in this case is simply the number of times they appeared in photographs with others. Eigenvector centrality takes into account how well connected a node's neighbours are. A node with relatively low degree can have high eigenvector

Figure 3.7: PKK Leaders: the 15 Most Central Nodes



The 15 most central nodes in the PICAN correspond to known leaders of the PKK. Different measures of centrality shed light on the nature of their image co-appearances, which correspond to functional and factional divisions within the PKK.

centrality if they are connected to well-connected nodes (Goldbeck, 2015). As such, eigenvector centrality is often interpreted as a measure of popularity in social networks (Borgatti et. al., 2013). Betweenness centrality measures the fraction of shortest paths that pass through a given node (Jia et. al. 2012). Nodes with high betweenness centrality occupy a powerful structural position in the network as they function as a "broker or a bridge" between subgroups and the rest of the network (Hansen et. al., 2020). Figure 3.7 displays the 15 most central nodes in the PICAN using all three measures, ordered by betweenness centrality.

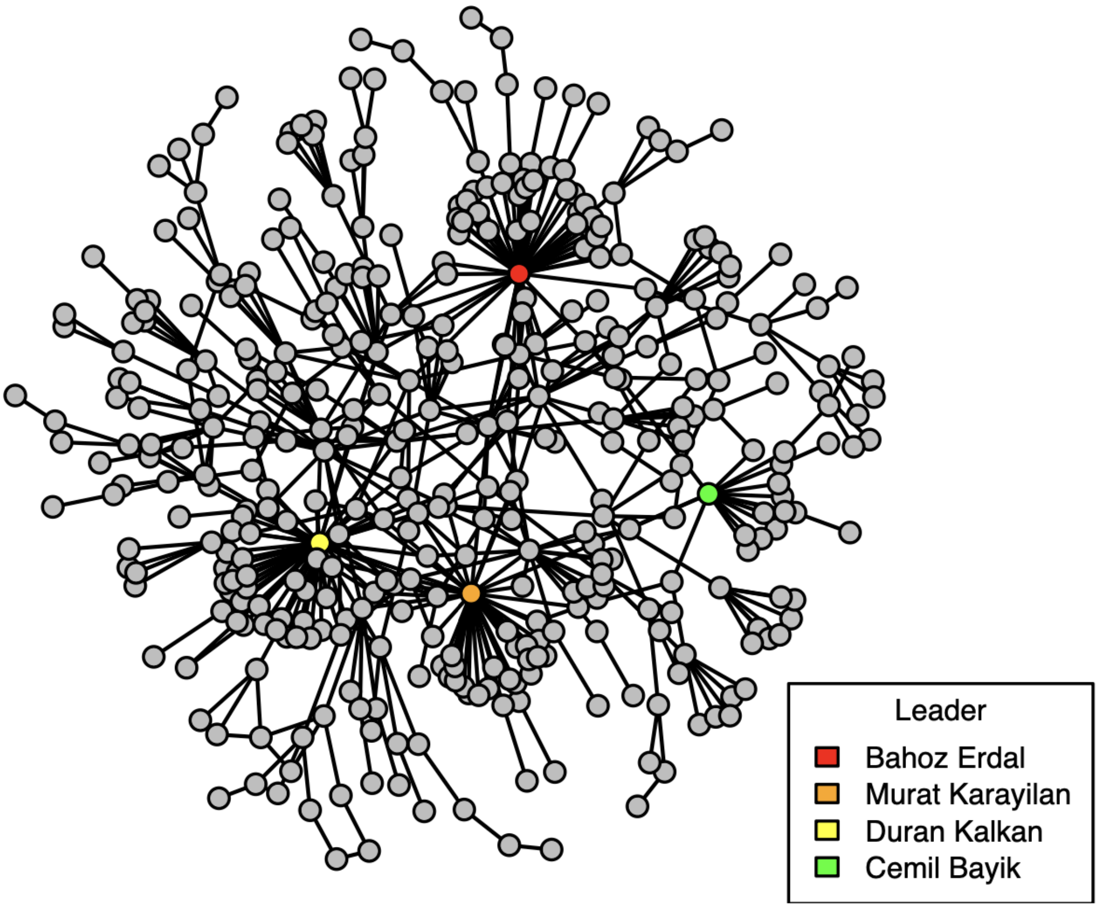
There are three nodes that stand out regardless of which measure of centrality is employed: Bahoz Erdal, Murat Karayilan, and Duran Kalkan. The node with the highest betweenness centrality in the PICAN is that of Bahoz Erdal, who has commanded the PKK's armed wing, the HPG (Hêzên Parastina Gel) since its founding in 2004. Despite his seniority in the PKK's military apparatus, he is not a member of the PKK's Executive Committee. Images of Erdal date back decades, and he appears to be primarily posing with rank and file militants; he is frequently the only connection for many of his neighbouring nodes. As the PKK's top military commander, his ubiquity in pictures taken in the field is to be expected. Though Karayilan has a slightly lower betweenness centrality than Erdal, his eigenvector centrality is more than twice as great. Karayilan is the current leader of the PKK, and frequently appears alongside other members of the PKK's executive council and top commanders as well as rank-and-file members. Unlike Erdal, many of his neighbours are also high-degree nodes rather than rank-and-file members, inflating his eigenvector centrality. The same is true of Duran Kalkan, who is also a member of the Executive Committee. Intuitively, those who appear most frequently with others tend to be high-ranking individuals. However, a closer examination of those that do not have many co-appearances despite their high rank demonstrates that internal political discord is organically reflected in these photographs.

3.5.1 Leadership Struggles

A particularly interesting node in this regard is that of Cemil Bayik, one of the founding members of the PKK and one of the three current members of the PKK's Executive Committee alongside Karayilan and Kalkan (Stansfield and Shareef, 2017). He was considered the PKK's "longtime number two" after founder Abdullah Öcalan, and was the top commander of the PKK's former armed wing, the ARGK (Artêşa Rizgariya Gêle Kurdistan) (Gunter, 2000). Despite being a central figure in the PKK's development, Bayik's node in the PKK network is far less central than those of his peers in the Executive Committee, Karayilan and Kalkan. His eigenvector centrality is particularly low– the lowest in the sample above, which is surprising given that the other Executive Committee members have the highest eigenvector centrality in the network by virtue of frequently posing with other high ranking individuals. Indeed, in this sample of nearly 20,000 photographs spanning decades, Bayik– one of the highest ranking members of the PKK on paper– does not appear in a single photograph with any of the other top leaders of the organization, including the two other members of the Executive Committee.

This discrepancy appears to be the result of a failed leadership struggle that occurred during the PKK's reorganization following its military defeat in 1999. The PKK dissolved the military wing led by Bayik (the ARGK), and replaced it with a new wing called the HPG, led by Bahoz Erdal. Indeed, "when Bahoz Erdal became the head of the PKK's military wing in June 2004, Bayik apparently found himself

Figure 3.8: PKK Leadership in the PICAN LCC



The position of four senior members of the PKK are shown in the PICAN. Karayilan, Kalkan, and Bayik are the three members of the PKK's Executive Committee. Erdal is the leader of the PKK's armed wing, the HPG.

in a lesser position." (Gunter, 2014: 65). The reasons for Bayik's ouster were likely related to his fraught relationship with the PKK's founder; "Öcalan has criticized Bayik's military leadership and ability to direct a guerrilla movement. According to Öcalan, Bayik prefers to stay behind the frontlines and has become involved in several controversial cases. For instance in 1992, fearing capture by the Turkish security forces, Bayik killed 17 wounded PKK militants in a cave." (Uslu, 2008). When the PKK resumed military operations in 2004, Bayik was consigned to diplomatic duties and spent much of his time in Iran, far from the battlefield. As a result, Bayik simply appears in far fewer photographs than his peers in the Executive Council.

Ali Haydar Kaytan is another individual whose position in the PICAN is relatively marginal despite being a co-founder of the organization, ranking last in the top 15 nodes in terms of betweenness centrality. Much like Bayik, Kaytan was also the subject of an internal purge and was sent by the PKK to fulfill diplomatic duties in Europe (Çandar, 2020). The relatively low centrality of Bayik's and Kaytan's nodes in the PICAN effectively captures their marginalization within the organization, despite their early importance to the group's development, and Bayik's membership of the Executive Committee.

3.5.2 Functional Divisions

Differences between eigenvector and betweenness centrality largely align with divisions between the PKK's political and military wings. Nodes with a greater betweenness centrality function as bridges between rank-and-file members and the leadership, and generally occupy combat roles. Nodes with greater eigenvector centrality primarily appear with high ranking individuals rather than rank-and-file members, and tend to be individuals in non-combat positions. Bahoz Erdal, the PKK's top military commander, displays greater betweenness centrality than eigenvector centrality. Gürhan, Welat, and Akdogan display a similar pattern of co-appearance to Erdal. All three joined the PKK at a young age as recruits, garnered significant battlefield experience, and rose through the ranks eventually becoming military commanders (Sidar, 2020; ANF News 2018; hpgsehit.com, 2022). Most of their co-appearances are with individuals who have few if any other co-appearances, likely subordinate cadres. Celik and Kaytan display greater eigenvector centrality than betweenness centrality. Though Celik was a military commander, he was also a member of the PKK's political assembly (ANF News, 2013). Kaytan has largely been involved in non-combat roles including leadership of the KCK, a pan-Kurdish political organization that subsumes the PKK (Sabah, 2020).

People who frequently appear with others in these photographs tend to be high ranking. However, the nature of these co-appearances– whether one primarily appears with foot soldiers, high-ranking individuals, or both– even enables a rudimentary distinction between members of the political and military wings of the organization using different measures of centrality. Indeed, the strength of the relationship between how socially embedded an individual is (i.e. their centrality in the network) and their influence within the PKK is evidenced by the fact that the PICAN even captures the effect of internal political struggles; the discrepancy between Cemil Bayik's status as a member of the PKK's Executive Committee and his relatively low centrality in the PICAN reflects his marginalization following a failed leadership struggle. Rather than being a source of bias, the performative nature of photography in the context of militant photographs is itself a rich source of contextual information that can be interpreted through an examination of the nature of the co-appearances.

3.6 Node Centrality and Rank

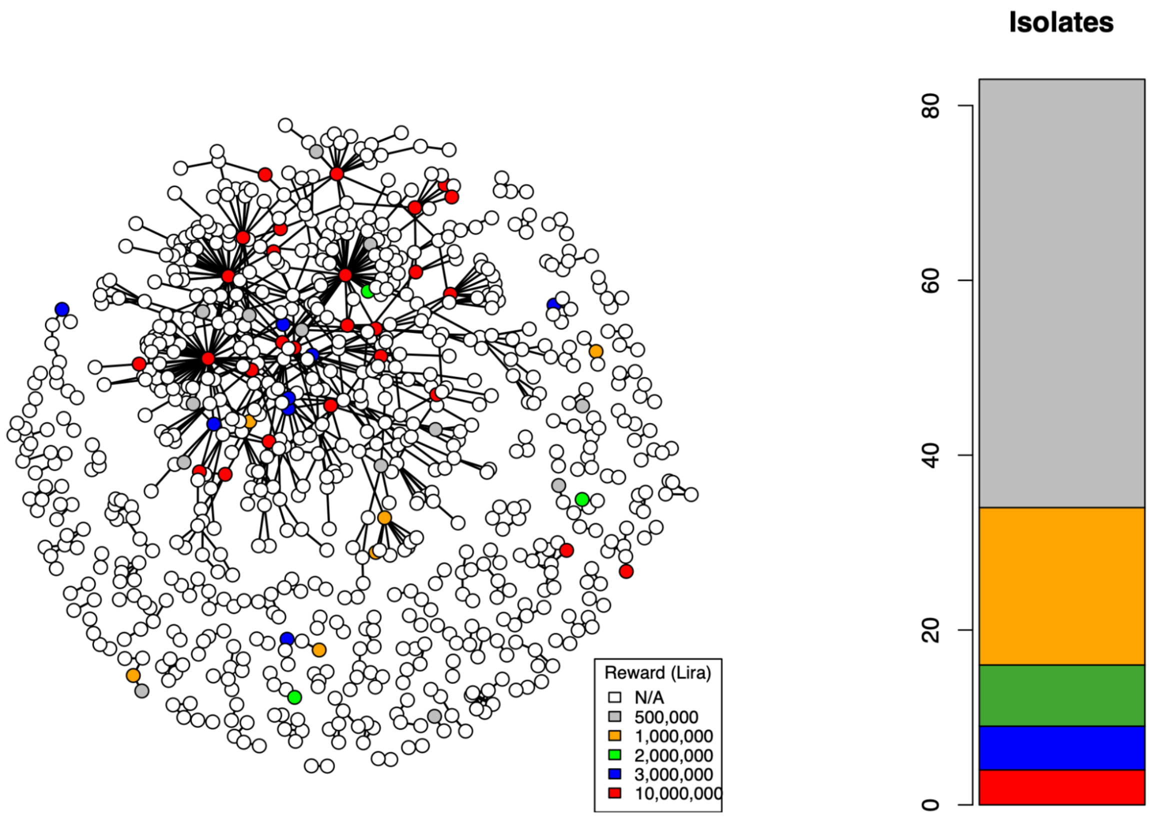
The previous section established that image co-appearance can highlight the top leadership of a rebel group, and even shed light on political divisions therein. However, the link between photogenicity and rank may only exist for the highest ranking (and thus, most salient) members of an insurgent group: while new recruits may clamor for a photograph with the supreme leader, a platoon commander may inspire less enthusiasm. This section matches 126 individuals from wanted lists maintained by the Turkish government with nodes in the PICAN to test whether there is a relationship between an individual's rank within the PKK (as measured by the reward offered for their capture) and their corresponding node's position in the PICAN. Linear regression models suggest that over half of the variation in the monetary reward offered for an individual can be explained by image co-appearances alone. Exponential Random Graph Models show that higher ranking individuals are significantly more likely to form ties in the PICAN. These results are not simply driven by the top leadership posing with new recruits; In both sets of models, the relationship between seniority and centrality is robust to the exclusion of the 28 highest-ranking officials (those on the Red List), suggesting that photogenicity gradually increases with rank.

The Turkish government maintains lists of wanted persons (including those killed, captured, and at large), offering monetary rewards for information leading to their capture. Individuals are assigned into one of five color coded wanted lists (red, blue, green, orange, and grey), with reward amounts ranging from 10 million Turkish

Lira for individuals on the Red List to 500,000 TL for individuals on the Grey List. The value of the reward offered for an individual's capture (i.e. the color of the wanted list that they are assigned to) is largely a function of the Turkish government's assessment of that individual's rank within the PKK. Article 6 (2) of the "Regulation on Rewards to be Given to Those Who Help in Discovering Terrorist Crimes" states that individuals are assigned to one of these five lists by "[...] grouping them according to their hierarchical position in the terrorist organization and/or the weight of the consequences of their actions, and by specifying the maximum amount of reward that can be given to those in each group" (Turkish Ministry of Interior, 2022). This legislation, available on the wanted website, further specifies that the highest reward amounts are offered for individuals on the Red List, who are deemed to be "senior manager[s] in a terrorist organization". Though no precise information is given regarding how exactly individuals are assigned to different lists, individuals on the Blue List seem to represent the second echelon judging by the relatively small number of individuals in this category (49), as well as the reward amount (3 million). The Green and Orange Lists likely correspond to PKK members who are believed to have some level of managerial responsibility that justifies offering rewards of 2 million and 1 million, respectively. The Grey List offers the lowest reward amount and is by far the largest group, with over 500 listed individuals, which may suggest that they are known to be members of the PKK, but do not occupy positions of power.

There are some important limitations that complicate the relationship between reward value and rank. In particular, certain low-ranking individuals may be put in a higher wanted category because they are deemed to be an imminent threat to public safety; for example, a footsoldier who was identified as having been involved in a recent attack may be put on the Red List. The relatively large number of individuals on the red list could suggest that this is the case. In a conventional military, one would generally expect to observe an inverse relationship between the level of seniority and the number of people at that rank; 82% of the U.S. military consists of enlisted personnel, 18% are officers, and just 1.25% occupy the highest rank (CRS, 2010). In the wanted list, the number of individuals in each category generally decreases as the reward value rises; there are 581 individuals on the Grey List, 158 on the Orange List,

Figure 3.9: PICAN Nodes Matched with Wanted Lists



In the graph of connected components on the left, nodes in the PICAN are matched with wanted persons and colored according to the value of the reward offered for their capture. The bar on the right counts the number of isolates that were matched at each wanted level. Individuals on the Grey and Orange Lists– the least wanted categories– tend to be isolates. Individuals on the higher-priority Red, Blue, and Green lists tend to be well connected.

73 on the Green List, and 49 on the Blue List. There are, however, 178 people on the Red List which may suggest that for some individuals the high reward is not a function of their rank but of their perceived threat to public safety. Alternatively, this could also reflect selection bias in intelligence gathering, whereby resources are directed towards identifying the top leadership of the PKK rather than trying to identify mid-level officers. Despite these limitations, the wanted list provides a credible external reference of an individual's rank within the PKK.

To examine the relationship between network centrality measures and rank, individuals from the wanted lists were matched with nodes in the network using facial recognition and text-based matching. Each entry on the wanted lists contains a photograph of the individual, their full name, date and place of birth, and the organization they are alleged to be a member of. Facial recognition was performed using the individual's photo on the wanted website, employing the same euclidean distance threshold (0.39) as the one used for in the clustering step of the construction of the PICAN network. Text-based matching was also conducted by identifying exact matches in terms of first name and last name between both datasets. Both facial and text-based matches were verified manually, yielding a total of 126 confirmed matches.

Figure 3.9 shows the position of nodes that were matched with records on the wanted list, coloring them according to their wanted level. The graph of connected components on the left shows that individuals on the Red and Blue lists occupy highly central positions in the network. The bar chart on the right suggests that very few of these individuals are isolates. Conversely, individuals on the Grey and Orange lists are largely absent from the graph of connected components or occupy relatively peripheral positions. The bulk of these individuals have no co-appearances at all.

Table 3.1: Summary Statistics: Matching PICAN Nodes with Government Watchlists

		Red	Blue	Green	Orange	Grey
Panel A:	Reward (TL, 000)	10000	3000	2000	1000	500
Matching Statistics	Listed Individuals	178	49	73	158	581
	Matched	28	9	6	20	59
	% Matched	15.7	18.4	8.2	12.7	10.2
Panel B:	Image Count	14.17	12.17	3.0	7.67	6.95
Network Centrality	% Isolates	13.79	58.33	71.43	85.71	84.48
	Degree	0.0235	0.0068	0.0014	0.0014	0.001
	Eigenvector	0.0629	0.0332	0.0102	0.0008	0.0071
	Betweenness	0.0571	0.0222	0.0	0.0017	0.0015

The columns in this table indicate the colors of five wanted persons lists maintained by the Turkish Government. Panel A reports the value of the reward offered for individuals on each list, the number of listed individuals, and the number and proportion of PICAN nodes that were matched with these individuals. Panel B reports attributes for matched nodes in each list, including the average number of images they appear in, the proportion that are isolates, as well as three measures of centrality.

The number and relative proportions of individuals identified in each wanted category are reported in Panel A of Table 3.1. The generally low match rates can be attributed to the inherent challenges of identifying members of a clandestine organization using facial recognition and imagery from two different sources. In particular, poor image quality posed a significant challenge for facial recognition– most of the images on the wanted list website were unusable due to being stretched to fit the dimensions of the image element, and only one image was available for each individual.

Failed matching due to image quality generates data that is Missing Completely At Random (MCAR), which simply reduces the sample size. However, differing match rates between wanted lists raises concerns about selection bias. The highest match rates were among those in the most wanted categories– 16% of the individuals on the Red List and 18% of those on the Blue List were matched to nodes in the PICAN. Lower match rates were observed for the Green, Orange, and Grey lists, which could indicate that lower-ranking individuals are underrepresented in the obituary images. However, the lowest-ranking group (the Grey List) has the largest total number of matches, over twice as many as the group with the next-highest number of matches.

Panel B of Table 3.1 reports the average values of several network statistics for each group. On average, individuals on the Red List appear in 14 photos, while individuals on the Grey List appear in just 7. The proportion of nodes in each group that are isolates decreases as the reward amount increases; only 14% of nodes in the Red List were isolates, increasing to 84% in the Grey List. The same is true of all three measures of centrality. The summary statistics and the graph in Figure 3.9 seem to suggest a positive relationship between node centrality and an individual's rank within the PKK as determined by the value of the reward placed on them by the Turkish government.

3.6.1 Linear Models

To further explore the relationship between node centrality and rank, I specify a simple OLS regression model of the following form:

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Where Y is the reward amount (in thousands of Turkish Lira) offered for individual i, and x is a measure of node centrality. Because measures of centrality are highly correlated with each other, a separate regression is run for each metric. Three common measures of centrality are used: degree, eigenvector, and betweenness. These are standardized by the number of images an individual appears in to emphasize social embeddedness rather than photogenicity without introducing multicollinearity between an image count variable and the centrality measures. For simple degree centrality, this effectively yields a measure of co-appearances per image. Table 3.2 summarizes the results of these regressions.

Table 3.2: OLS Regression Coefficients for Standardized Network Centrality Measures and Reward Value

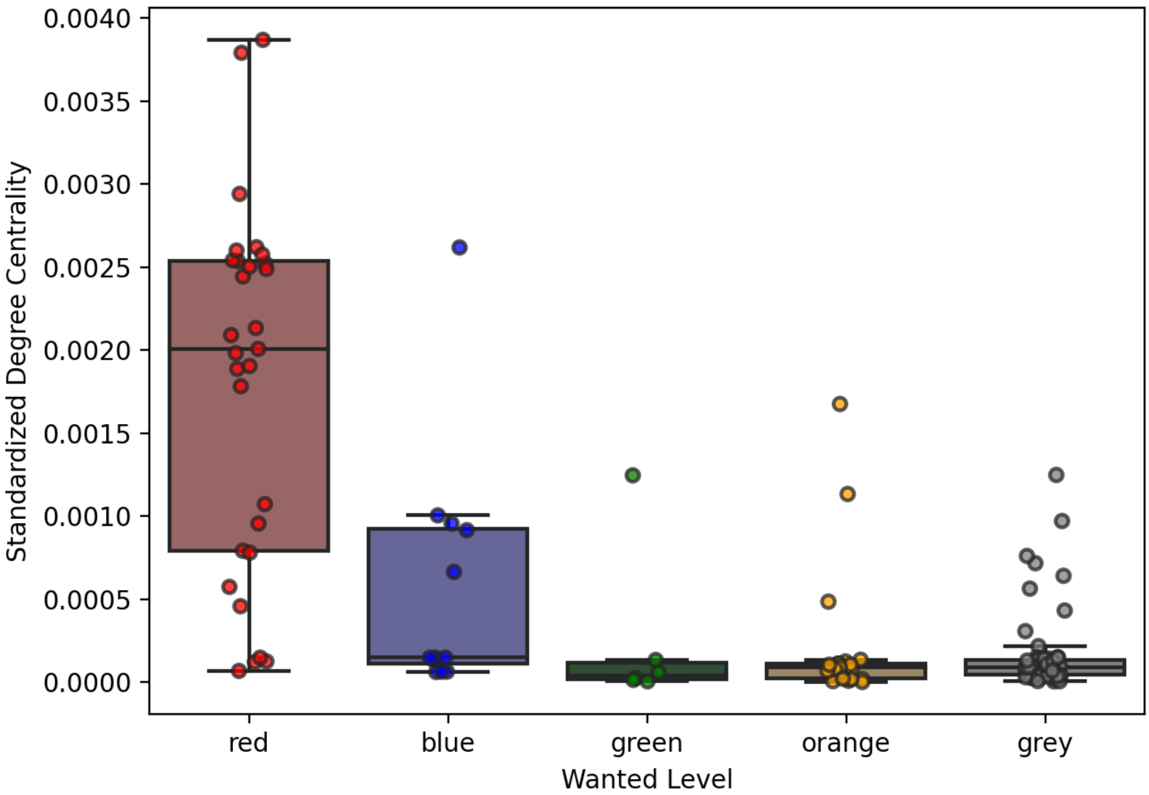
	Model 1	Model 2	Model 3	Model 4
(Intercept)	2051.32***	1466.35***	2442.66***	1950.80***
	(445.61)	(269.17)	(339.06)	(297.39)
Image Count	113.63***			
	(33.00)			
Degree		3134293.53***		
		(258269.53)		
Eigenvector			352246.85***	
			(69378.62)	
Betweenness				1117667.87***
				(124487.94)
R^2	0.09	0.54	0.17	0.39
Adj. R^2	0.08	0.54	0.17	0.39
Num. obs.	126	126	126	126

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

The dependent variable in all models is the value of the reward placed on an individual by the Turkish Government. The independent variable in Model 1 is the number of photographs an individual appears in. Models 2, 3, and 4 use different measures of node centrality divided by the number of images that individual appears in.

Model 1 regresses the count of images an individual is identified in with the value of the reward placed on them. Though not a measure of centrality per se, this provides useful contextual information on whether higher ranking individuals appear in more photographs than those of lower rank. Indeed, the results from the first column indicate that for each additional photograph an individual appears in, the reward amount increases by 114,000 Turkish Lira on average. However, less than 10% of the variation in reward amounts is explained by the number of photos an individual appears in. Models 2, 3, and 4 show strong positive associations between degree, eigenvector, and betweenness centrality and reward values. The model with the greatest explanatory power relies on simple degree centrality, achieving an R^2 value of 0.54; in other words, over half of the variation in the bounty price placed on an individual can be explained by the frequency with which they appear with others in photographs. The boxplot in Figure 3.10 shows the relationship between standardized degree centrality and wanted level.

Figure 3.10: Network Centrality and Wanted Level



The colors represent wanted lists, which offer monetary rewards for an individual's capture. They are arranged in descending order of reward value: 10m, 3m, 2m, 1m, and 0.5m. The Y axis represents degree centrality divided by the number of images an individual appears in. Individuals on the red and blue lists tend to be far more central than those on the lower-priority lists.

To ensure that these results are not entirely driven by extreme values, the models in Table 3.2 are re-run excluding the 28 individuals who appear on the Red List. The top leaders of rebel groups are often well known, but information on the structure of rebel groups below the level of top leadership is much harder to come by; second and third order commanders are less prolific. By excluding the top leadership from the regressions, the results in Table 3.3 assess whether node centrality measures not only distinguish the top echelon from the rank-and-file, but whether they can distinguish lower ranks as well.

Model 1 shows a substantial decrease in the effect size and significance of the relationship between the number of photos an individual appears in and their reward value when the red list is excluded. This reflects some of the idiosyncrasies of picture-taking in the PKK: foot soldiers can have a relatively large number of portrait photos,

Table 3.3: OLS Regression Coefficients for Network Centrality Measures and Reward Value, Excluding Red List

	Model 1	Model 2	Model 3	Model 4
(Intercept)	672.40***	913.71***	932.82***	901.13***
	(153.90)	(86.44)	(87.19)	(80.83)
Image Count	44.96*			
	(17.22)			
Degree		53811.91**		
		(16026.47)		
Eigenvector			8802.63**	
			(3108.03)	
Betweenness				27185.59***
				(5722.21)
R^2	0.07	0.11	0.08	0.19
Adj. R^2	0.06	0.10	0.07	0.18
Num. obs.	97	97	97	97

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

The models in this table are identical to those presented in Table 3.2, except that the 28 individuals in the most wanted category (the Red List) are excluded. Results are robust to this exclusion.

but higher ranking individuals (particularly those that are still alive) rarely appear alone. Despite reductions in effect size, significance, and model fit across the board, Models 2-4 nonetheless display a consistent relationship between node centrality and wanted level even when the highest echelon is removed from the analysis.

3.6.2 Exponential Random Graph Models

Though Borgatti et al. (2013) note that hypothesis testing using node-level attributes is mostly conducted via standard regression models in the literature, the use of network data violates the assumption that observations are independent. This problem is overcome by Exponential-family Random Graph Models (ERGMs), which enable an assessment of the statistical likelihood of observing certain network configurations, including the influence of nodal attributes on tie formation. ERGMs treat the observed network (in this case, the PICAN) as one realization from a set of possible networks with similar core characteristics (Robins et. al. 2007). The general form specifies the probability of the observed network as a function of a set of network features that may occur more or less frequently than expected by chance:

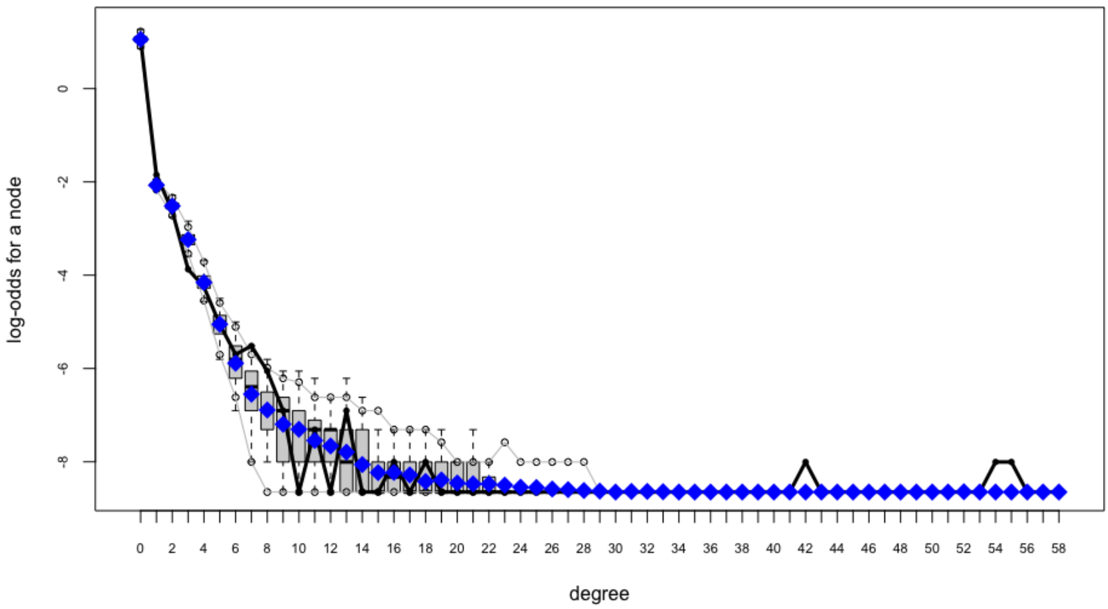
$$P(Y = y) = \frac{exp(\theta' g(y))}{k(\theta)}$$

Where Y is the random variable for the state of the network, and y is the realization thereof. The right hand side specifies a vector of coefficients θ for model statistics g of the network y, normalized by the sum of all possible networks $k(\theta)$ (Statnet, 2021). These networks are generated via Markov Chain Monte Carlo maximum likelihood estimation (MCMCMLE). If more or less of a certain network configuration is present in the observed network than in the simulated one, the parameter corresponding to that configuration is iteratively adjusted in subsequent simulations until the simulated network closely approximates the observed one (Ellwardt et. al. 2012). For example, given a parameter for the number of edges involving female nodes, a positive value would indicate that women have a greater propensity to form ties in the PICAN than in a random graph.

Table 3.4 reports the results from four ERG Models which assess whether high-ranking PKK members (i.e., those fetching higher bounties) are more likely to co-appear in images than lower ranking individuals. Models 1 and 2 utilize the full PICAN and include a parameter for isolates, while models 3 and 4 exclude isolates altogether. The inclusion of a triadic close term in the full PICAN leads to model degeneracy, so it is only included in the analysis of connected components.

All models converged successfully, indicating that the final simulated network in each case was sufficiently close to the observed network in terms of the included parameters. To test the extent to which the Exponential Random Graph Model presented in Table 3.4 fits the data, it is convention to observe how well the model reproduces network properties that are not included as parameters in the estimation. The main hypothesis being tested in the model is monadic in nature (whether the nodal attribute of reward value influences the likelihood of tie formation). As such, similar to Stys et. al. (2021), Figure 3.11 compares a node-level network statistic– the degree distribution– between the observed and simulated networks. The simulated networks largely reproduce network statistics in the PICAN that were not included as terms in the ERGMs, suggesting that the models faithfully approximate the observed network.

Figure 3.11: ERGM Goodness of Fit Diagnostics



The empirical distribution is represented as a solid black line, while the simulated distributions are indicated by the boxplots. Though three very high degree nodes– the PKK's main leaders– consistently defy simulation, the rest of the observed degree distribution is generally within the bounds of the simulated models.

The inclusion of network statistics enables accurate modeling of the observed network, and provides important contextual information. All models in Table 3.4 include the variable "Edges", which captures the baseline probability of edge formation. The negative coefficient suggests that the baseline probability of edge formation is lower than it would be in an equivalent random graph. This is partly driven by the high number of isolates, as shown by the roughly 20% reduction in the size of this coefficient when isolates are excluded. Models 1 and 2 include a separate parameter for isolates, which is positive and significant. The large number of isolates captures a specific type of photoshoot conducted by the PKK in which an individual poses for a large number of photographs in which they are the only subject (and thus have no co-appearances). This is particularly common among seemingly lower-ranked individuals. Models 3 and 4 include a term for triadic closure, which is positive and significant, indicating that edges are more likely the more closed triangles they create. The large number of triangles in the PICAN is a structural feature of an image co-appearance network: because all individuals in a group photo co-appear with each other, a group photo

Table 3.4: Effect of Reward Value on the Likelihood of Edge Formation: Parameter Estimates from ERGMs

	Model 1	Model 2	Model 3	Model 4
Edges	-9.69^{***}	-9.53^{***}	-7.90^{***}	-7.55^{***}
	(0.20)	(0.33)	(0.16)	(0.33)
Isolates	2.09^{***}	2.08^{***}		
	(0.08)	(0.08)		
Triadic Closure			1.72^{***}	1.71^{***}
			(0.06)	(0.06)
Age	0.04^{***}	0.04^{***}	0.02^{***}	0.02^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
Gender	0.13^{**}	0.12^{**}	-0.00	-0.01
	(0.05)	(0.05)	(0.05)	(0.05)
Reward	0.14^{***}		0.07^{***}	
	(0.01)		(0.01)	
Wanted (Unmatched)		-0.08		-0.19
		(0.14)		(0.15)
Wanted (1M)		0.04		0.09
		(0.23)		(0.24)
Wanted (2M)		-0.27		-0.73
		(0.48)		(0.52)
Wanted (3M)		0.59^{**}		0.24
		(0.19)		(0.19)
Wanted (10M)		1.30^{***}		0.51^{**}
		(0.16)		(0.16)
AIC	14129.42	14133.88	10476.49	10471.77
BIC	14196.00	14253.72	10529.52	10567.23
Log Likelihood	-7059.71	-7057.94	-5233.24	-5226.89

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Models 1 and 2 use the full PICAN and include a paramter for isolates. Models 3 and 4 exclude isolates, but include a parameter for triadic closure (Geometric Edgewise Shared Partners or GWESP). Age and Gender are node-level attributes. "Reward" is a continuous node-level attribute indicating the value of reward placed on an individual by the Turkish government. "Wanted" is a categorical version of this variable, with the reward for each category indicated in parentheses.

containing n people generates a clique with $3\binom{n}{3}$ triangles. Thus, a photo of three people generates three triangles, but a photo of five people generates 30 triangles[1]. The inclusion of this parameter led to degeneracy in the full models and was therefore only included in the analysis of connected components.

All models also include apparent age and gender as control variables. Gender has a modest but significant effect in the full network: women are 1.14 times more

[1] $3\binom{5}{3} = 3\frac{5!}{2!3!} = 3\frac{120}{2\times 6} = 3 \times 10 = 30$

likely to form ties than men, which is consistent with the prior observation that women appear in more images[2]. This likely reflects the performative nature of many of these photographs, which seek to emphasize the role of women in the PKK. Age also increases the likelihood of edge formation, which may result from the fact that older members have simply had more time to appear in pictures. Alternatively, age could also be proxying for seniority. Age and rank are likely to be correlated in any military setting: infantrymen are rarely old, generals are rarely young, and the PKK's known leaders are all above the age of 60. Despite this, the "Reward" variable remains positive and significant in both models, though the magnitude of the effect is diminished by the addition of controls.

The main variable of interest in all models is the value of the reward placed on an individual by the Turkish government. Models 1 and 3 include a continuous version of this variable, denoting the value of the reward offered for an individual in millions of Turkish Lira. In both the full network and the restricted sample, nodes corresponding to individuals with a higher reward value have a significantly higher probability of forming edges. In the full network, a node on the Red List is 10.92 times more likely to form a tie than a node on Grey List, ceteris paribus[3]. This effect is halved when isolates are excluded, likely due to the fact that 85% of the nodes on the Grey list are isolates.

Models 2 and 4 treat the reward variable as categorical rather than continuous. This enables the estimation of a separate effect for each reward amount and allows us to set an appropriate base category; though there are five different wanted lists corresponding to different reward amounts, this variable contains a sixth category: PICAN nodes that were not matched to wanted individuals. This constitutes an inappropriate reference group because information about these individuals' rank is unknown; as such, each level of the "Reward" variable is compared against individuals on the Grey list (0.5M TL). Individuals on the Red and Blue wanted lists (fetching rewards of 10M TL and 3M TL, respectively) are significantly more likely to co-appear in images than individuals on the Grey List in the full network. However, the coefficient

---

[2] $e^{0.13} = 1.14$

[3] $(10 \times e^{0.14}) - (0.5 \times e^{0.14}) = 10.92$

on the Blue List category becomes insignificant in Model 4, due to the exclusion of isolates (and thus over 85% of the observations in the Grey List reference category). Despite the categorical nature of the variable, the ratio between the coefficients and the reward amounts of Red and Blue list members is remarkably close: the reward for an individual on the Red List is triple that offered for someone on the Blue List, and in both models the coefficient of the Red List is roughly double that of the Blue List. This aligns with the OLS results in Table 3.2, suggesting that co-appearance not only identifies the top echelon of the PKK's command structure, but lower ranking officers as well. However, individuals on the Green and Orange lists (1M and 2M) did not have a significantly higher likelihood of co-appearance relative to the Grey list, though this is likely due to the small number of observations in these categories.

The absence of a significant relationship between membership of the Green and Orange lists and likelihood of edge formation suggest that the relationship between social embeddedness and node centrality is not consistent across the full range of the data. Indeed, the lower match rates between individuals on these two lists and nodes in the PICAN imply that some lower-ranking officers may be disproportionately absent from the obituary photos altogether. Nevertheless, results from this section demonstrate a general relationship between an individual's rank within the PKK (measured in terms of reward value) and the centrality of their corresponding node in the PICAN. Results from the linear regressions indicate that over half of the variation in reward values can be explained by the frequency with which an individual appears with others in photographs. All three of the centrality measures vastly outperform image count in terms of explaining variation in reward values, suggesting that social embeddedness is a better proxy for rank than photogenicity. ERGM results are consistent with these findings, with higher value individuals significantly more likely to form edges in the PICAN. Importantly, this relationship persists across all models even when the highest-ranking officials were excluded, meaning that results are not driven purely by the propensity of new recruits to pose with the group's leader.

An important caveat to this analysis is that the PICAN is only an approximation of the PKK's structure. Co-appearances in these images can reflect a number of forms of social interaction, from friends sharing a meal to a new recruit posing with their commander. Edges do not provide any direct information on the nature of the relationship between two individuals, other than the fact that they crossed paths. At their most basic level, edges in the PICAN simply reflect the number of social contacts an individual has had. In aggregate, the quantity and nature of these interactions correlates strongly with an individual's rank, allowing for the creation of a network that approximates the known structure of the PKK. It is important to note, however, that some individuals can be important to the organization without being highly socially embedded or visible in photographs from the field, and conversely that some highly photogenic individuals may be of low systemic importance.

3.7 Network Structure and Counterinsurgency Tactics

A significant gap between the theoretical and empirical literature on the application of SNA to the study of insurgent groups involves the relationship between network topology and robustness to counterinsurgency strategies (Helfstein and Wright, 2011). This section seeks to assess the extent to which a co-appearance network generated from militant photographs can help to fill this gap.

Despite the strong association between rank and centrality, the PICAN does not represent an organogram of the PKK; it is a representation of how the PKK chooses to present itself to the world. This exaggerates certain characteristics, such as the group's female cadres, while understating others such as lower-level commanders, the "dishonorably discharged", and the politically marginalized. Nevertheless, given a large enough sample it seems plausible that an image co-appearance network would capture the general structural characteristics of an insurgent group– the rough number and relative proportions of individuals at various ranks, and the relationships between them. In what follows, I analyse the topology of the PICAN in reference to historical developments in the conflict between the PKK and the Turkish government.

Ünal (2016) identifies two main phases in Turkish military strategies towards the PKK: ground war and targeted killings. Until the year 2000, the Turkish military largely waged a conventional ground war against the PKK. Under this doctrine, the

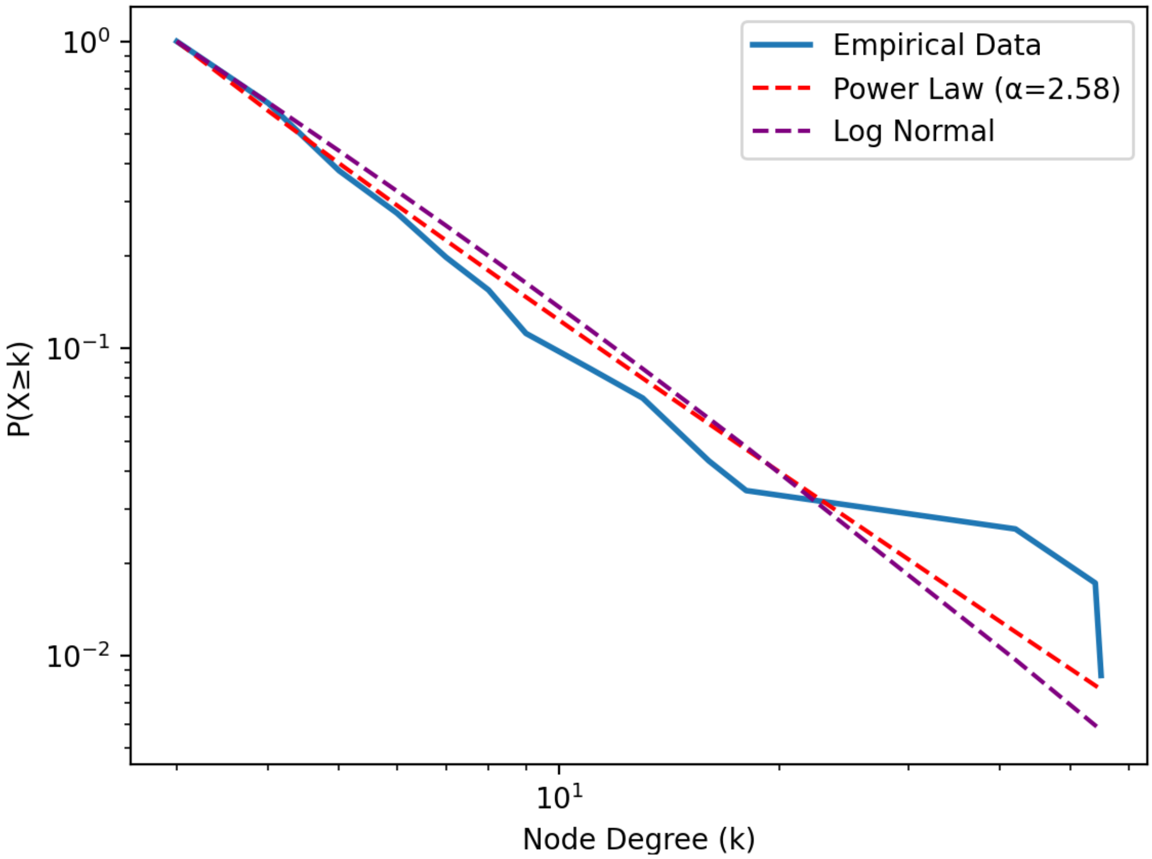
Turkish government conducted large scale military operations, armed roughly 60,000 civilians to act as paramilitary force and burned as many as 4,000 Kurdish villages to the ground (Gürcan, 2015; Filkins, 2003). Though this strategy culminated in the PKK declaring military defeat in 2000, it failed to prevent the group's resurgence only four years later. Following a period of low intensity conflict, large-scale hostilities resumed between the PKK and Turkish forces following the breakdown of a truce in 2015. This new phase in the conflict has been accompanied by a significant doctrinal shift on the part of the Turkish government towards a more targeted approach. Rather than relying on large military operations and poorly trained paramilitaries, the new approach relies heavily on the use of special forces, drone strikes, and intelligence gathering (Ünal, 2016). Particularly since 2020, the Turkish government's focus has been on targeting the leadership structure of the PKK (Jamestown, 2021).

These various counterinsurgency strategies can be simulated, and their effects on the network can be assessed empirically. In general, the static robustness of a network to intentional attacks "analyses the ability of a system to maintain its connectivity after the disconnection or deletion of a series of targeted nodes. In this context, the connectivity of the resulting network is typically measured by the size of the largest connected component (LCC)" (Lordan and Albareda-Sambola, 2019:1). This section proceeds with an empirical analysis of the two main counterinsurgency strategies in chronological order; first, the strategy of inflicting maximum casualties against the PKK through a ground war, and second, the doctrine of targeted strikes against high ranking members.

3.7.1 Ground War

Much of the historical literature concerning first phase of the conflict between the Turkish Government and the PKK focuses on a puzzling dynamic: despite an aggressive military doctrine focused on engaging the PKK in a conventional war resulting in the group declaring military defeat, the PKK not only survived, but was able to regroup with remarkable speed. "By the early 2000s, the PKK was militarily weak, and the majority of its members were outside Turkey's borders. However, after a few years of calm, on 1 June 2004, the PKK put an end to its unilateral ceasefire and once again

Figure 3.12: PICAN Scale-Free Degree Distribution



The PICAN's degree distribution (shown in blue) follows a power law of $k^{-2.58}$, making it a weakly scale-free network.

began to attack civilian and military targets in Turkey. How was the PKK able to survive and rebuild itself in such a short period of time?" (Pusane, 2015: 727). Thus, in the early stages of the conflict, the Turkish government tried to defeat the PKK by inflicting the maximum number of casualties possible, with little concern for whom within the organization was being targeted. Translating this policy into the realm of social networks, this doctrine roughly equates to a strategy of random node removal.

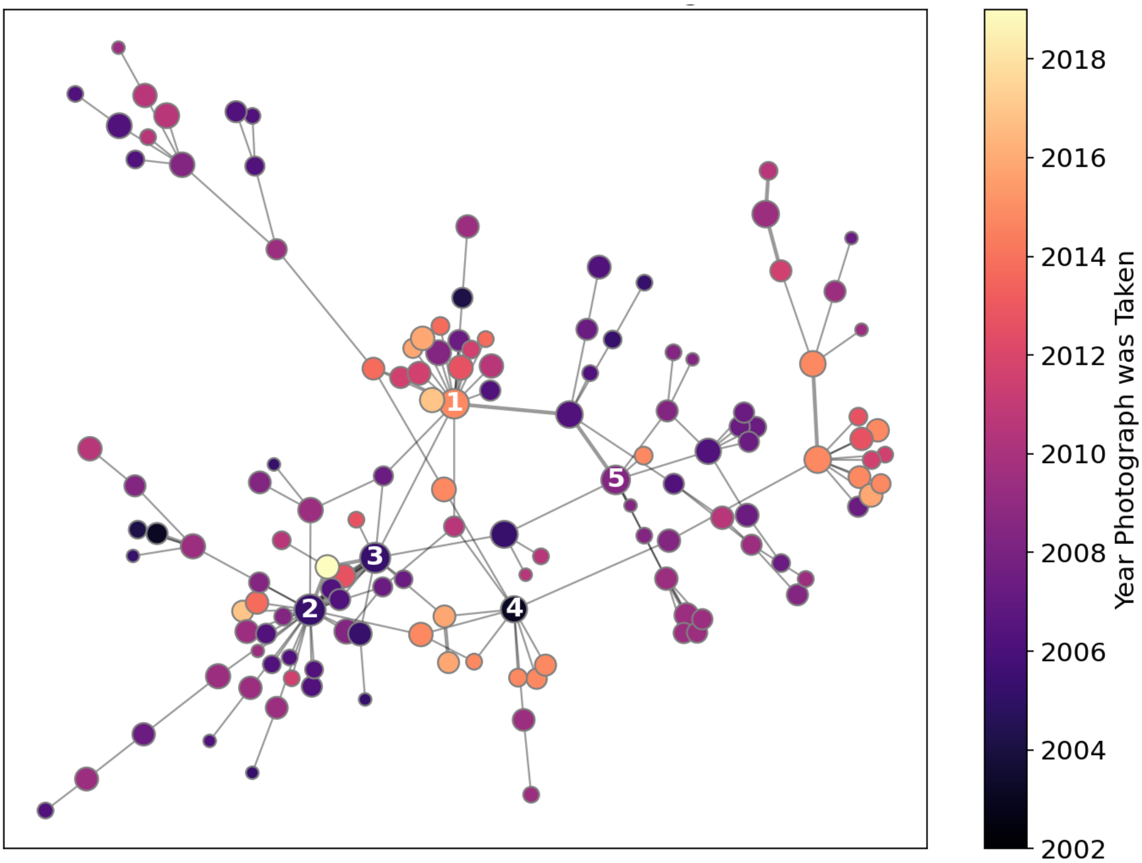
The robustness of a network to various types of node removal depends largely on its structure. A salient characteristic that can be observed in the PICAN is its highly uneven degree distribution. As shown by the previous section, despite the majority of nodes in the network being either isolates or low-degree, a small number of nodes have very high degree. A scale free network is a graph in which the nodes' degree distribution follows a power law, where the probability $p(k)$ that a node has degree k is a function of $k^{-\alpha}$, such that $p(k) \propto k^{-\alpha}$, where α is a constant. Figure 3.12 shows the empirical distribution of node degrees, compared with a power law distribution.

The empirical degree distribution appears to loosely follow a power law distribution of $k^{-2.58}$. According to Bodio and Clauset (2019), "a network is deemed scale free if the fraction of nodes with degree k follows a power-law distribution $k^{-\alpha}$, where $\alpha > 1$", making the PICAN weakly scale-free. These types of networks have properties that are of direct relevance to counterinsurgency tactics; in particular, "a compelling characteristic of scale-free networks is that a significant fraction of nodes can be removed randomly without the network losing connectivity" (Verma and Zhang, 2020). This is because the vast majority of nodes have a low degree, while only a small number of nodes are responsible for the overall connectivity of the network. If nodes are removed at random, it is highly unlikely that a high-degree node would be removed.

A counterinsurgency strategy based on inflicting the maximum amount of casualties against the PKK is analogous to a strategy of random node removal. This was the strategy pursued by the Turkish Armed Forces (TSK) in the 1990s, and it succeeded in weakening the PKK militarily to the point that "the PKK was defeated conclusively in 1999 and several scholars argued that the party would never recover" (Plakoudas, 2018). However, despite the fact that the PKK incurred heavy losses, these were mostly foot soldiers. The leadership structure was left intact, allowing the organization to regroup after only four years.

Following its military defeat in 2000 and re-emergence in 2004, the PKK grew rapidly. Figure 3.13 utilizes image metadata to track this growth, showing the evolution of the network over time. Timestamps from photographs are extracted, and the earliest timestamp for each node is chosen thereby indicating an individual's first known appearance in the field. Of the images that were sorted into clusters, 41% (6373) contained machine-readable timestamps. This analysis should be taken as an approximation of the PKK's regrowth rather than a precise model thereof. Many images in the sample do not contain readable timestamps, meaning that an individual may have been present in the PKK for far longer than the length of time suggested by the metadata; For example, the earliest available timestamp for Duran Kalkan is in 2014, despite his cluster containing several photos that appear far older.

Figure 3.13: Preferential Attachment Over Time using EXIF Data



A network is constructed from metadata-containing images. Nodes are colored according to the oldest timestamp in that node's image cluster. The five most central nodes are numbered, and correspond to leaders. Newer nodes tend to cluster around older nodes.

The top 5 most central nodes (in terms of betweenness) in this restricted sample are labeled, and once again correspond to the PKK leadership. Node 1 is Duran Kalkan, who despite the light color of his node has been in leadership roles within the PKK since the 1990s. The bulk of his neighbours appeared between 2008 and 2014. Nodes 2 and 3 are Bahoz Erdal and Kadir Celik, two senior figures in the military wing of the PKK, the HPG. They co-appear most frequently with individuals who first appear in the early 2000s, though their sustained presence in the field is indicated by co-appearances with much more recent recruits as well. Node 4 is Ali Haydar Kaytan, a co-founder of the PKK who largely served in non-combat roles. Interestingly, he appears almost exclusively with fairly recent recruits who first appear from 2014 onwards. Node 5 is Mehmet Gurhan, a military commander whose neighbours first appeared in the early 2000s. Despite incurring heavy losses throughout the 1990s, the PKK was able to regroup around a small number of core members.

Figure 3.13 shows some evidence of preferential attachment, whereby new nodes attach preferentially to already well-connected nodes. As newer, lighter nodes appear, they tend to attach to darker and more well-connected nodes. In other words, as newer recruits begin to appear in photographs, they tend to appear alongside older, more established members. Importantly, it seems that different elements of the leadership were important hubs at different periods in the PKK's resurgence. This likely reflects the fact that "the PKK developed an advanced organizational structure, which included both military units and those elements responsible for recruitment activities, ideological training, propaganda efforts in order to increase awareness about the Kurdish question, and fundraising." (Pusane, 2015: 728). Erdal, Celik, and Gurhan, members of the military leadership, were important hubs in the immediate aftermath of the PKK's return to armed struggle in 2004. Political leaders, such as Kaytan, Kalkan, and Karaylian, act as important hubs for more recent recruits, particularly those appearing since 2014. The PKK's ability to withstand generalized attacks (i.e. random node removal), as well as its regrowth around key members (preferential attachment) are characteristics associated with scale free networks. Thus, the structure of the PICAN aligns closely with the historical resilience of the PKK itself.

3.7.2 Targeted Strikes

Assuming the topology of a co-appearance network generated from militant photographs loosely approximates the general structural characteristics of the group itself, the former can be used to make general inferences about the future success of more recent counterinsurgency strategies.

Seemingly in recognition of the failure of its previous counterinsurgency strategy, the Turkish government has embarked on a radically different approach in recent years. Particularly since 2016, Turkish forces have focused on targeting the leadership structure of the PKK through the use of airstrikes and raids both within and beyond Turkey's borders, enabled by growing domestic production of drones such as the Bayraktar TB-2 (Tastekin, 2016). In 2016, Turkish forces erroneously claimed to have assassinated Bahoz Erdal, the current leader of the PKK's armed wing. The claim was spread largely by pro-government media, and indicated that he had survived at least two prior assassination attempts (Sabah, 2016). In 2021, two senior PKK officials– Ali Haydar Kaytan and Sofi Nurettin– were killed in airstrikes in Northern Iraq (Sabah, 2021). Rather than attempting to inflict the largest possible number of casualties against the PKK by engaging in a conventional ground war, Turkish forces appear to be focusing on targeting key individuals within the organization.

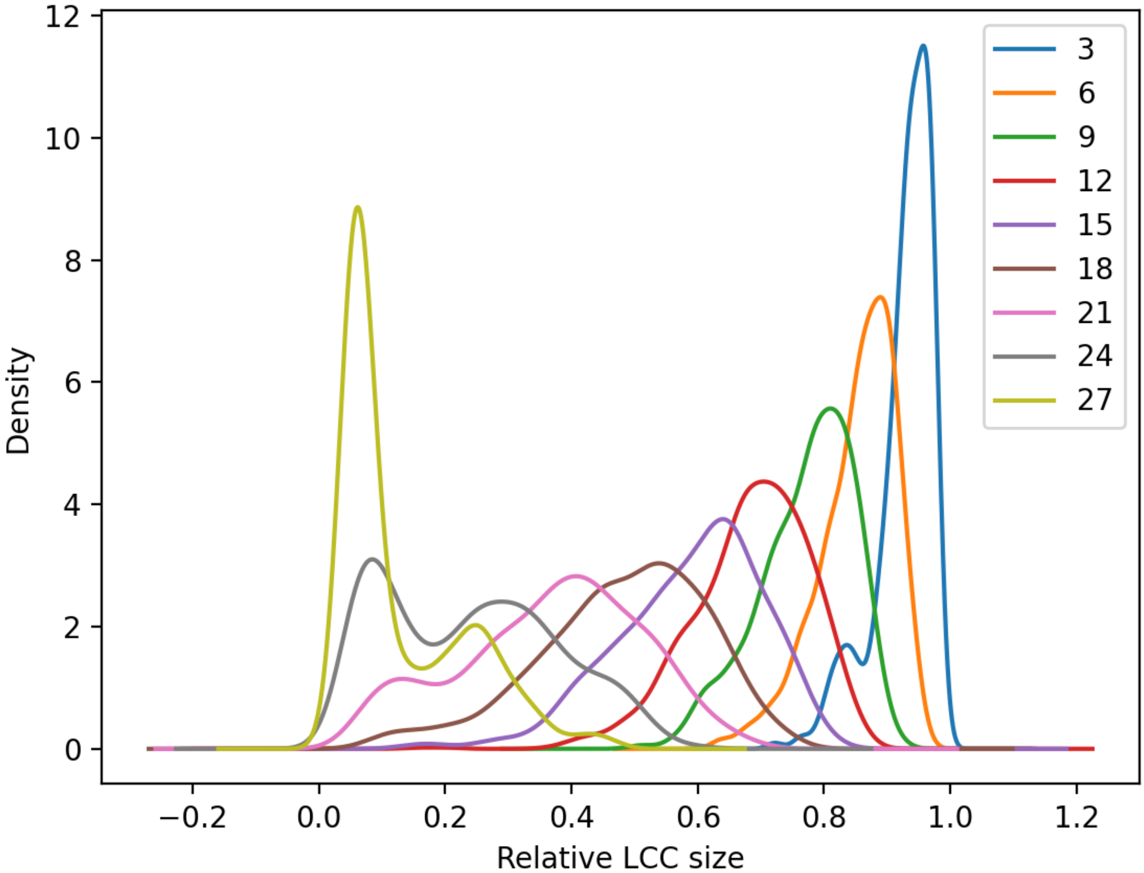
In theory, a counterinsurgency strategy predicated on targeting the senior leadership of the PKK should be more disruptive to the network than one based on the removal of nodes at random; given the scale-free nature of the network, "disabling just a few critical nodes can result in a disconnected network especially for the smaller nodes" (Verma and Zhang, 2020). An empirical assessment of this claim in reference to the PKK fills an important gap in the applied literature on the social network analysis of insurgent groups: "claims that scale-free militant networks are resilient due to their robustness against random attacks on nodes have not adequately reckoned with the countervailing effect that high-degree nodes are more visible" (Zech and Gabbay, 2016: 233).

The selection of targets in this strategy is likely largely opportunistic; previous operations have targeted PKK leaders on the basis of available intelligence, which often results from chance encounters. In May 2021, the Turkish President announced the assassination of Sofi Nurettin, a PKK military officer responsible for the group's presence in Syria (Sabah, 2021). The operation was carried out by an intelligence unit formed specifically to target Nurettin, which ultimately located him by tracking his family (Sabah, 2021). Nurettin survived an assassination attempt in 2016, suggesting that Turkish forces had been targeting him for at least 5 years (Al-Monitor, 2016). Beyond entailing a significant investment in terms of military and intelligence resources, the operation's success hinged significantly on luck. Thus, the Turkish government's new military strategy involves targeting PKK leadership, but only managing to strike opportunistically.

The distributions shown in Figure 3.14 approximate the effects of such a strategy on the LCC by randomly removing a given number of nodes T from the set of 30 most central nodes. Each level of T roughly corresponds to the aggressivity of a counterinsurgency strategy; the removal of only 3 random nodes out of the top 30 most central nodes would correspond to a relatively weak counterinsurgency approach, while the removal of 27 out of 30 corresponds to an extremely aggressive strategy. In practice, the value of T is likely to be determined by available opportunities. 1000 random sets of nodes are removed at each level, and the resulting kernel density estimates of the effect on LCC size are reported.

If only three out of the top 30 most central nodes are removed (the scenario indicated by the distribution in blue), there is virtually no effect on network connectivity. As more nodes are removed, connectivity decreases but the variance of the distributions increases substantially. Thus, the extent to which the removal of, for example, 15 nodes decreases the size of the LCC depends heavily on which 15 nodes are removed. Bimodality in some of these distributions likely indicates splintering brought about by either the removal or survival of the three most central nodes. When 27 of the top 30 most central nodes are removed, there is a strong likelihood of the relative LCC size dropping to near zero. However, the smaller second peak on the right of the distribution indicates that the LCC can retain roughly 20%– and in rare cases over 40%– of its original size, depending on which three nodes out of the top 30 survive. Similarly, though the removal of only 3 of the top 30 nodes overwhelmingly results in the LCC remaining unchanged, a much smaller peak to the left of the distribution

Figure 3.14: PICAN Robustness to Opportunistic Central Node Removal

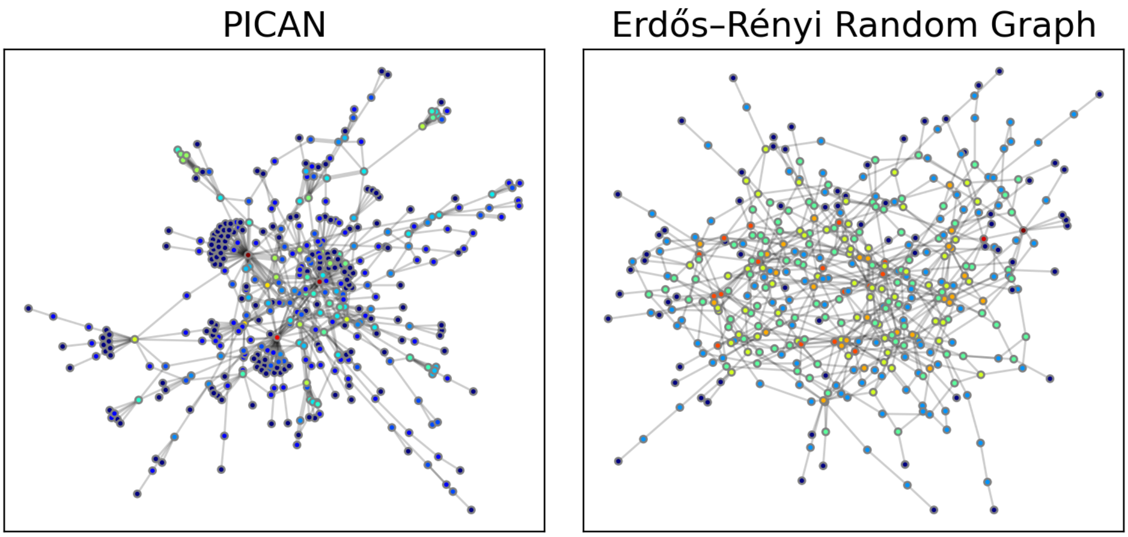


Distributions reflect the effect of removing a random set of nodes from the top 30 most central nodes on the relative size of the LCC. Colors indicate the number of nodes removed, a procedure which is repeated 1000 times. High levels of connectivity are maintained even when the majority of the top 30 nodes are removed.

indicates potential reductions in LCC size of up to 20%, presumably in the unlikely event that the three most central nodes are removed.

Though the theoretical literature on scale-free networks suggests that these are highly vulnerable to the removal of hub nodes, the PICAN appears to be surprisingly robust to such a strategy. Even when a random set of 15 of the top 30 most central nodes is removed, the network tends to maintain well over 60% of its connectivity. Thus, while the top three most central nodes in the PICAN stand out in terms of their centrality, a significant amount of the network's connectivity is distributed across a larger set of important figures. Indeed, the FARC– a similarly hierarchical paramilitary organization– was able to withstand the assassination of 53 of its leaders and three members of its Secretariat (Segura and Mechoulan, 2017).

Figure 3.15: Comparison between PICAN and a Random Graph



Compared to a random graph with the same number of nodes and edges, the PICAN features a number of high-degree hub nodes (shown in red). There also appears to be a higher degree of clustering and interconnectivity in the PICAN.

This type of resilience is characteristic of "small world" networks. This class of networks overlaps with (and shares many of the properties of) scale-free networks; they are characterized by the presence of "hub" nodes, and are thus generally robust to random node removal. Small world networks are further characterized by a short average path length; compared to a random graph with the same number of nodes and edges, the presence of well-connected hubs enables shorter paths between sets of nodes. Figure 3.15 compares the Largest Connected Component of the PICAN with an Erdős–Rényi random graph with the same number of nodes and edges, the former displaying the aforementioned hub-and-spoke configuration.

In the PICAN, the average shortest path between nodes is 4.9; the random graph has an average shortest path length of 6.4. This is largely due to the presence of high-degree hub nodes that are clearly visible in the PICAN but absent in the random graph. The nodes are colored by degree, with the random graph exhibiting a much more uniform degree distribution. Importantly, the PICAN not only contains well-connected hub nodes, but general areas of high connectivity. This type of interconnectivity is also typical of small world networks, which exhibit a higher degree of clustering– the extent to which a node's neighbours are also interconnected. A quantitative measure of the extent of clustering in a graph is given by the Watts and Strogatz (1998) clustering coefficient, defined as:

$$c_i^{ws} = \frac{2E_i}{k_i(k_i - 1)}$$

Where E_i is the number of edges between neighbours of node i, and K is the degree of the node. The more interconnected a node's neighbours are, the higher the clustering coefficient, and the more robust that network will be to the removal of the node. The average clustering coefficient of the PICAN is 0.0574 compared to just 0.0014 for the random graph, meaning that the former is more densely interconnected and therefore robust to the removal of even high-degree nodes.

Humphries and Gurney (2008) propose a formal test for small-world-ness (3) based on the premises that, compared to a random graph, (1) a small-world network will have a shorter average path length L and (2) a higher degree of clustering C:

$$\lambda_g = \frac{L_g}{L_{rand}} \tag{3.1}$$

$$\gamma_g^{ws} = \frac{C_g^{ws}}{C_{rand}^{ws}} \tag{3.2}$$

$$S^{ws} = \frac{\gamma_g^{ws}}{\lambda_g} \tag{3.3}$$

If the above conditions are met, then $S > 1$ and the network may be considered a small world. The value of S for the PICAN is 2.69, suggesting that the organizational structure of the PKK resembles that of a small world network.

In a study of jihadist insurgent groups, Sageman (2011: 140) notes that "Unlike a hierarchical network that can be eliminated through decapitation of its leadership, a small-world network resists fragmentation because of its dense interconnectivity". Jemaah Islamiyah, a Southeast Asian jihadist group, was effectively incapacitated following the arrest of its senior leadership; the group's extremely hierarchical nature left local cells unable to function without direct orders from above (2011: 141). Despite the presence of a few key hub nodes in the PICAN, there also exists a large pool of highly interconnected nodes representing the second and third tiers of the organization's leadership. This affords the PKK a high level of robustness to the targeting of even its most central figures. Furthermore, the node removal simulations carried out in Figure 3.14 are carried out on a static graph. As shown in the analysis of the PKK's regrowth shown in Figure 3.13, the network's evolution over time is characterized by preferential attachment to hub nodes. Not only would the survival of even a small number of these nodes allow for the regrowth of the group; "new hubs will take the role of the eliminated ones and restore the network's ability to function." (Sageman, 2011: 141).

3.7.3 Limits of the PICAN

Quantitative estimates in this context must be taken as illustrative rather than definite. As previously mentioned, the PICAN only approximates the structure of the PKK based on measuring the social embeddedness of individuals using image co-appearances. Similarly, though measuring the relative size of the LCC following various node removal strategies provides a simplified way of understanding the group's structural resilience to different counterinsurgency strategies, it does not provide a direct measure thereof. Indeed, there are a number of important disconnects between the PICAN and the actual structure of the PKK brought about by the performative nature of insurgent photography that affect the conclusions that can be drawn from this analysis.

Firstly, certain individuals may fulfill critical functions within an insurgent group that are not captured by their centrality in an image co-appearance network. The assassination of Sofi Nurettin, who was responsible for the PKK's presence in Syria, decreases the size of the LCC by just 1.17%. His connectivity in the PICAN almost certainly underestimates his importance to the PKK; Liaising with Syrian Kurdish YPG fighters may offer massive benefits in terms logistical and tactical coordination, but may not offer as many opportunities for photoshoots as being a platoon commander.

Secondly, new recruits clamoring for a photograph with the supreme leader collapses the chain of command, directly linking the lowest and highest ranking individuals in the PICAN. This creates a large number of "stars" in the network, where leaders (hubs) are connected to many foot soldiers (spokes). In actuality, this relationship would be mediated through a series of lower ranking officers, who are themselves somewhat under-represented in the PICAN according to the results from the previous section. The overall density of the network– as well as its resilience to targeted strikes– is likely underestimated: the removal of a hub disconnects large numbers of spokes, but if the spokes are more densely integrated into the network through layers of intermediary nodes, the removal of a hub has less of an effect.

Nevertheless, this section suggests that the robustness of a co-appearance network generated solely from militant photographs to certain node removal scenarios broadly aligns with the true network's historical resilience to analogous counterinsurgency strategies. The scale-free structure of the PICAN renders random node removal strategies ineffective, as connectivity is maintained by the survival of a relatively small number of high-degree nodes. This is congruous with the historical record: "Considering the fact that the PKK has developed such a complex system of networks and institutions, as well as various sources of funding both at the domestic and international levels, it was no surprise that Turkey's ability in weakening the PKK militarily in the late 1990s did not bring an end to the PKK insurgency" (Pusane, 2015: 730). Assuming even a tenuous relationship between the PICAN and the actual structure of the PKK, the image co-appearance network is able to provide contextual information relevant to an assessment of the future success of the counterinsurgency strategy currently being undertaken by the Turkish government. As a small world network, the PICAN tends to survive the removal of the majority of its most central nodes due to dense interconnectivity among them. Indeed, the preceding paragraph suggests that the PICAN probably underestimates the extent of this interconnectivity. The corresponding real-world conclusion is that the degree of hierarchy within the PKK is not so extreme as to enable the group to be incapacitated by the removal of a few key leaders. Thus, even a strategy of targeted strikes is unlikely to be effective against the PKK, especially when one considers the high cost and low success rate of such efforts.

Finally, the use of euphemistic terms such as “node removal” should not distract from the fact that this type of research is popular among those who seek to inflict harm. The military and intelligence communities are some of the primary consumers and patrons of research on social network analysis, particularly in the study of insurgency (Knoke, 2013). Marc Sageman (2011) a former CIA officer turned scholar, notes that the application of network analytic methods by the military and intelligence communities has largely been a failure because individuals in these organizations are tasked with applying network analysis to identify who to kill, and institutionally precluded from thinking about whether doing so will be effective (Bohannon, 2009). In contrast, the scholarly endeavor to understand insurgent groups as networked organizations has emphasized that the vast majority are exceptionally robust to precisely this sort of attack. Indeed, while the CIA used a social network analysis of call logs to identify and assassinate Osama Bin Laden (Knoke, 2013), Medina’s (2014) social network analysis of the global jihadist movement found that the assassination of both Bin Laden and Zarqawi did little to weaken the movement, as ”removing two hubs from the network does not severely damage the network structure and with respect to some metrics, does no damage at all.” In the absence of scholarly research tasked with understanding militant organizations as complex social and political entities that are hard to destroy but that can be engaged with productively, the status quo will persist. In the context of this paper, the Turkish government is already targeting the PKK’s top leadership; the ultimate conclusion from the empirical analysis herein is that such a strategy would be ineffective, and that conflicts with structurally similar groups such as the FARC have only been resolved through negotiated settlement.

3.8 Conclusion

The analytical approach elaborated in this study makes three distinct contributions to the literature on the application of Social Network Analysis to the study of insurgent groups. Firstly, it leverages an abundant but underutilized source of primary data generated by militants themselves. While the main challenge in the study of dark networks is the lack of data thereon, a defining trend in the evolution of modern insurgency has been the use of online media, particularly online visual propaganda (Dauber, 2020). Social network analysts have recognized the rich relational information contained in image co-appearances as a valuable tool for understanding social structures (Lewis et. al. 2008, Berry, 2006; Golder, 2008). By automating the process of generating an image co-appearance network through facial recognition and unsupervised clustering, the present approach is able to harness the vast quantities of unstructured image data generated by insurgent groups. The distribution of this methodology as an open source python package further facilitates its use by researchers of dark networks.

The ability to process large quantities of primary data enables a second significant contribution to the existing literature, which relies almost exclusively on secondary sources. Social networks generated through media reports, legal filings, and public statements require a significant amount of pre-existing knowledge about the group in question, and inherently omit all but the highest-ranking figures. Though selection bias in insurgent photographs is still a problem, a foot soldier is probably more likely to appear in the background of a group photograph than to be named in a newspaper. Image co-appearance networks are thus likely to be far more complete, and can be created without additional knowledge about the group in question.

Finally, the use of co-appearances imposes a level of consistency in the ordering of the network. In previous studies of insurgent groups, ties between nodes have encompassed everything from marriage to casual acquaintance (Gill et. al., 2014), from cohabitation to "weak contact" (Koschade, 2006: 567). To be sure, co-appearance in militant photographs encompasses a wide range of possible social interactions, from friends sharing a meal to a commander posing with her subordinates. If two people are co-present in a photograph, the nature and social significance of the link between them is unknown. However, if the same person repeatedly appears with others in a large number of photographs, there is probably an overarching reason. An image co-appearance network mirrors the social forces that govern when and where pictures are taken. In the context of an insurgency where virtually all social interaction is mediated by the group's structure, an image co-appearance network will generally reflect these ordering principles, including divisions between military units, ranks, and political factions: foot soldiers will co-appear because they're in the same platoon; commanders officiating training academy graduations will appear once with each graduand and frequently with each other; leaders who have fallen out are unlikely to pose together. The network will also reflect performative aspects of the group's self-perception, emphasizing actors and narratives that align with the image militants wish to present to the world, the precise nature of which depends on the ideological characteristics of the group in question.

Thus, the use of co-appearances in militant photographs enables detailed insights into the social, political, and organizational forces that structure insurgent groups. The three analytical sections in this study demonstrate the types of analyses that are possible with a co-appearance network generated by militant photographs, using a large quantity of obituary images taken by the PKK.

Section 5 demonstrates that a qualitative analysis of the nature of an individual's co-appearances can yield information on functional and factional divisions within a rebel group. In the PKK, military commanders have high betweenness centrality due to frequent co-appearances with subordinates, who are themselves generally poorly connected nodes. Political leaders spend less time with the rank and file, and more time with each other, leading to higher eigenvector centrality. The nature of co-appearances even captures the effects of Cemil Bayik's marginalization following a failed leadership struggle: despite being a founding member of the PKK, he has relatively few co-appearances and appears in no photographs with the top leaders of the organization.

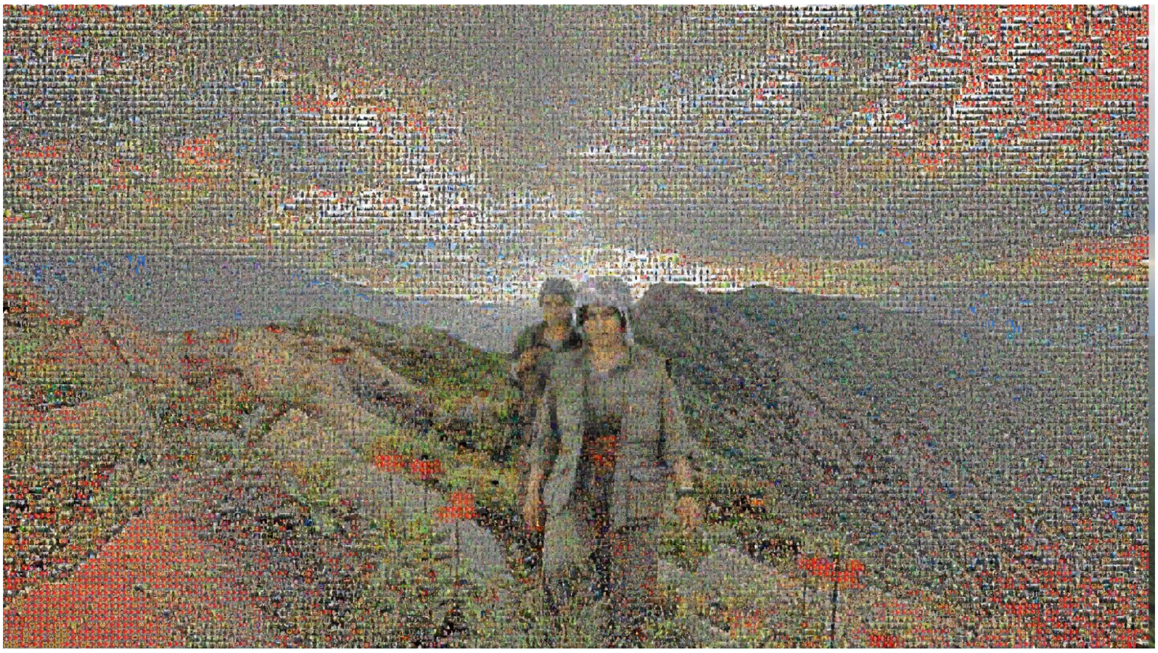
Section 6 demonstrates a consistent relationship between an individual's rank and their centrality in the co-appearance network. This is achieved by matching over 100 nodes in the PICAN with individuals appearing on lists of wanted persons maintained by the Turkish government, with the value of the reward placed on them acting as a proxy for rank. Linear models indicate that over half of the variation in the reward offered for an individual's capture can be explained purely by the number of times they appear with others in photographs. Exponential Random Graph Models show strong positive associations between reward value and the likelihood of edge formation, suggesting that higher ranking individuals are more well-connected in the PICAN. In both sets of models, results are robust to the exclusion of the top leadership from the analysis.

Section 7 shows that the co-appearance network may even approximate the general structural characteristics of the insurgent group. A close reading of the topology of the PICAN in reference to qualitative work on the PKK's history finds agreement between the PKK's resilience to generalized attacks and the PICAN's structural robustness to an analogous node removal strategy. If the PICAN can be assumed to roughly approximate the structure of the PKK, then the Turkish military's current approach of targeted strikes against leadership is also unlikely to be effective; a simulation of targeted node removal shows that the network's connectivity is maintained by a large and densely interconnected pool of high ranking individuals waiting in the wings.

While insights from these three sections are specific to the nature of the image dataset and the insurgent group in question, images gathered from other sources or rebel groups would manifest their own idiosyncrasies and highlight different social and organizational processes. For example, an egocentric network of a single insurgent cell or unit could be constructed from an individual's posts on social media, with temporal and geographical nodal attributes derived from metadata. Footage from the commission of war crimes– disturbingly rife online– could be used to understand whether incidents are occurring at random, or whether the same individuals systematically co-appear. Images collected from multiple groups could reveal the existence and nature of links between members of both organizations. Though dark groups are exceptionally difficult to study, the extraction of relational information from images in the form of co-appearances between individuals can yield detailed information on their social structure.

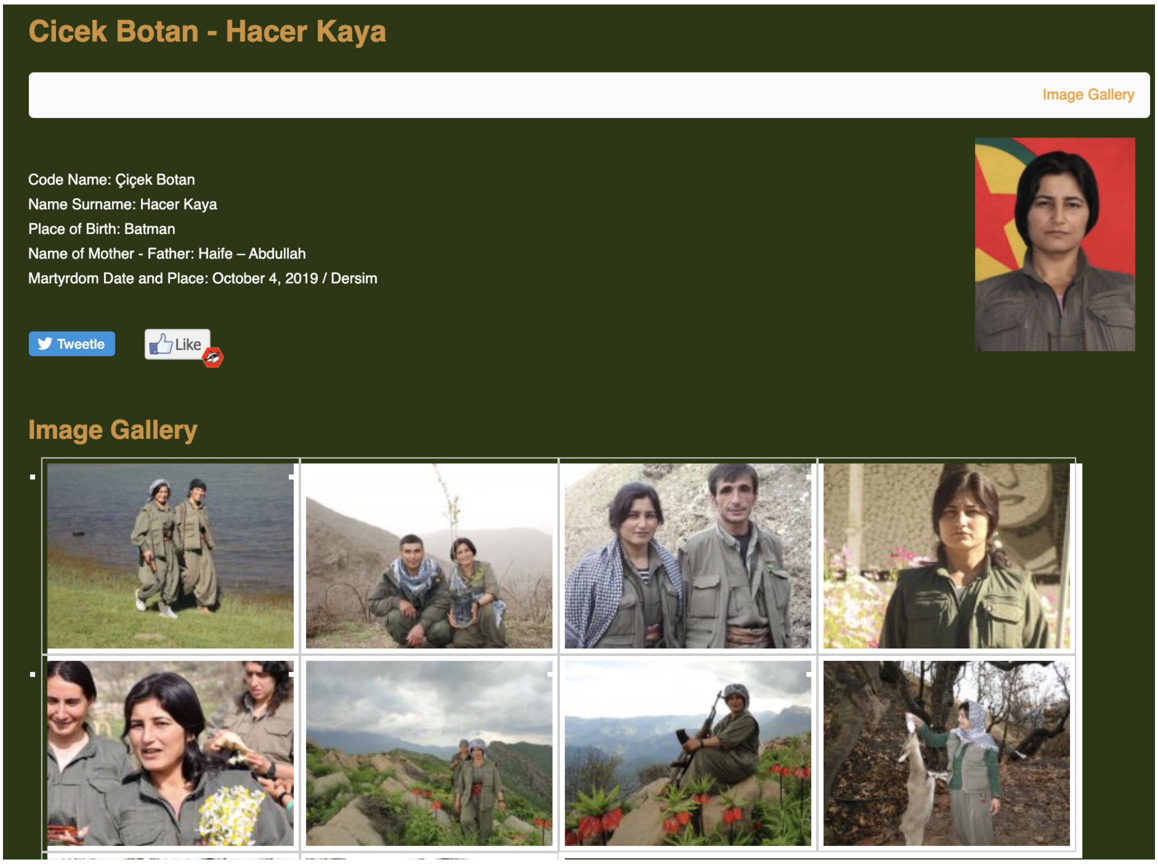
Appendix for Chapter 3

Figure C1: Mosaic of Input Images scraped from online PKK obituaries



This image mosaic was constructed from roughly 20,000 publicly available images posted by the PKK to their online obituaries website, hpgsehit.com. A version of this figure with an interactive zoom functionality to view individual pictures can be accessed here.

Figure C2: Sample Obituary



The screenshot above was taken from an entry on the PKK's obituary website. The obituary includes a high definition portrait in the top right, some basic biographical details on the left, and a set of images of the fallen militant.

Figure C3: Representations of Martyred PKK Leaders

The images above are samples from facial clusters containing representations of two PKK leaders who are no longer in the field, but whose likeness is still present in the photographs. On the right, the likeness of Mahzum Korkmaz is depicted on lapel pins which appear in 21 photographs. He was the PKK's first military commander, and despite being killed in 1986, remains an important figure in the PKK. The images on the left were taken from the cluster generated for PKK founder Abdullah Öcalan, who was arrested and imprisoned in 1999. His likeness is depicted on flags and murals that appear in the background of 107 photographs. Interestingly, despite not being in the field for the past 23 years, his node in the PICAN is the most central. Both of these phenomena reinforce the notion that there is a relationship between an individual's importance to the PKK and their embedding in the image co-appearance network. However, for the sake of conceptual clarity in the definition of edges in the PICAN, these nodes are removed from the network.

Figure C4: Central University Research Ethics Committee Approval Letter

Oxford Department of International Development
Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK
Tel: +44(0)1865 281800
qeh@qeh.ox.ac.uk www.qeh.ox.ac.uk

Head of Department: Professor Diego Sánchez-Ancochea

OXFORD DEPARTMENT OF INTERNATIONAL DEVELOPMENT

UNIVERSITY OF OXFORD

5 July 2022

Ollie Ballinger
ODID and St Anne's College

– Dear Ollie

Research Ethics Approval
Ref No: SSH/ODID DREC: C1A_22_047
Title: Rebellion as Complex Network: Social ties and hierarchy in the PKK

The above application has been considered on behalf of the Social Sciences and Humanities Interdivisional Research Ethics Committee (IDREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the Oxford Department of International Development's DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

Any data collection involving in-person interactions with participants must include an up-to-date fieldwork risk assessment. Please refer to the guidance at https://researchsupport.admin.ox.ac.uk/governance/ethics/coronavirus as the University's position on conducting in-person research may change. Please ensure you comply with any local COVID-19 restrictions and requirements.

Should there be any subsequent changes to the project that raise ethical issues not covered in the original application you should submit details to the DREC for consideration.

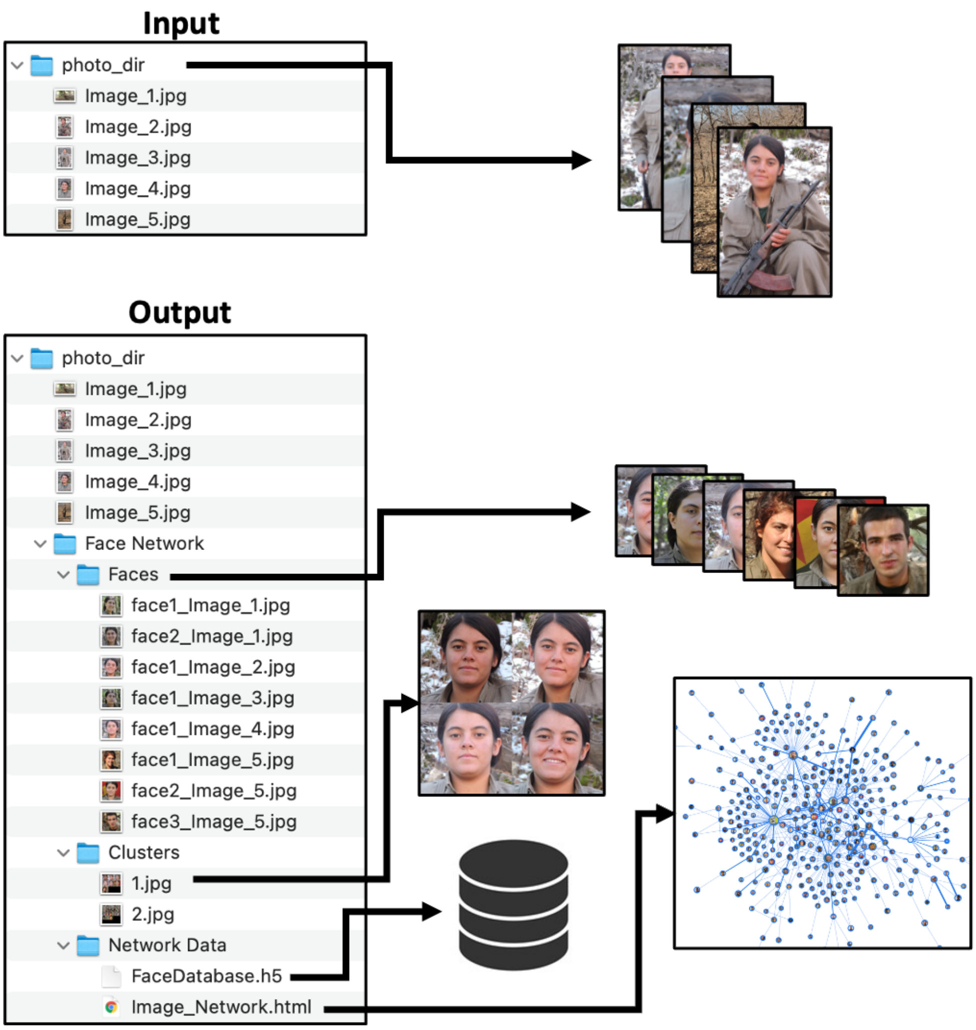
Yours sincerely,

Prof Loren Landau
ODID DREC Chair

The Face-Network Python Package

One of the main contributions of the third chapter is the development and release of an open source software package that allows researchers to apply the methodology elaborated herein to automatically generate social network graphs on the basis of unstructured image data. To this end, I have published the Face-Network package on the Python Package Index (PyPi). Using three functions, users can extract faces from photographs, identify individuals across photographs using clustering, and generate a network.

Given a folder of images "photo_dir", the execution of the three main functions in the face-network package would create a project directory with the following structure:



The code block below provides an example usage of the face-network package:

```
import face_network
photo_dir="Users/oballinger/Downloads/photo_dir"

face_network.extract(photo_dir, age_gender=True)

face_network.cluster(photo_dir, algorithm='DBSCAN', iterations=10, initial_eps=0.44, max_distance=40)

face_network.network(photo_dir, size=20)
```

Below, I provide documentation for the three main functions in the face-network python package, alongside performance metrics and explanations of the outputs from the functions.

1. Extracting Faces from Images

This function extracts all faces from a directory of images using Dlib's face detector, and must be run prior to further analysis.

```
face_network.extract_faces(source_dir, age_gender=False)
```

Parameters

source_dir:	(*str*); The path to the image folder (note: this folder can contain sub-directories so specify the highest level image directory).
age_gender:	(*bool default=False*); Estimates apparent age and gender using a pretrained Convolutional Neural Network (ResNet-50). Results are stored in the columns "age" and "gender" in the resulting dataframe. Gender is predicted on a scale from 0 (male) to 1 (female).

Outputs

This function creates a new folder called "Face Network" in your image directory. When a face is identified, it is cropped and stored in a new folder "source_dir/Face Network/Faces/". Given "Image.jpg" containing two faces, this function will save two cropped faces: "face1_Image.jpg" and "face2_Image.jpg". Facial encodings (128-dimensional vectors used for clustering and matching similar faces) are stored in a file called "FaceDatabase.h5".

Performance

The facial extraction function is threaded, by default using n_cpus-1. Using seven cores on an M1 MacBook Pro, all 14,000 faces were extracted from the ”Labeled Faces in the Wild” (LFW) database in 42 minutes (6 faces per second). Facial encodings for the entire LFW database take up 15 MB (1kb per face).

2. Clustering Similar Faces

Once faces are extracted, similar faces are clustered together. This function uses a density-based clustering algorithm (DBSCAN) to identify clusters of similar faces in the list of facial encodings. Starting with loose clustering parameters, the function iteratively decreases the neighborhood distance parameter. In each iteration, facial similarity within clusters is evaluated. Dense clusters are extracted, and sparse clusters are assigned to be re-evaluated in the next iteration. When an iteration returns no new clusters, the function returns a dataframe containing facial encodings grouped into clusters based on similarity.

```
face_network.network(source_dir, algorithm='GRAPH', iterations=1,
    initial_eps=0.45, max_distance=45)
```

Parameters

source_dir:	(*str*) The path to the image folder
algorithm:	(*{’GRAPH’, ’DBSCAN’, ’OPTICS’, ’AHC’}, default=”GRAPH”*) The algorithm used for clustering, with the default being the Chinese Whispers graph clustering algorithm used in this paper. Possible options are “DBSCAN”, “OPTICS”, and “AHC” (agglomerative hierarchical clustering), which may be better suited to different datasets.
iterations:	(*int, default=10*) The number of iterations that the function will perform. Each iteration restricts the clustering parameters.
max_distance:	(*float, default=50*) Sets the maximum euclidean distance between each face in a cluster and the core sample. This weeds out outliers and sparse clusters.
initial_eps:	(*float, default=0.45*) Sets the Epsilon parameter for the DBSCAN algorithm.
mosaic:	(*bool, default=True*) Creates a mosaic of face tiles for each image.

Outputs

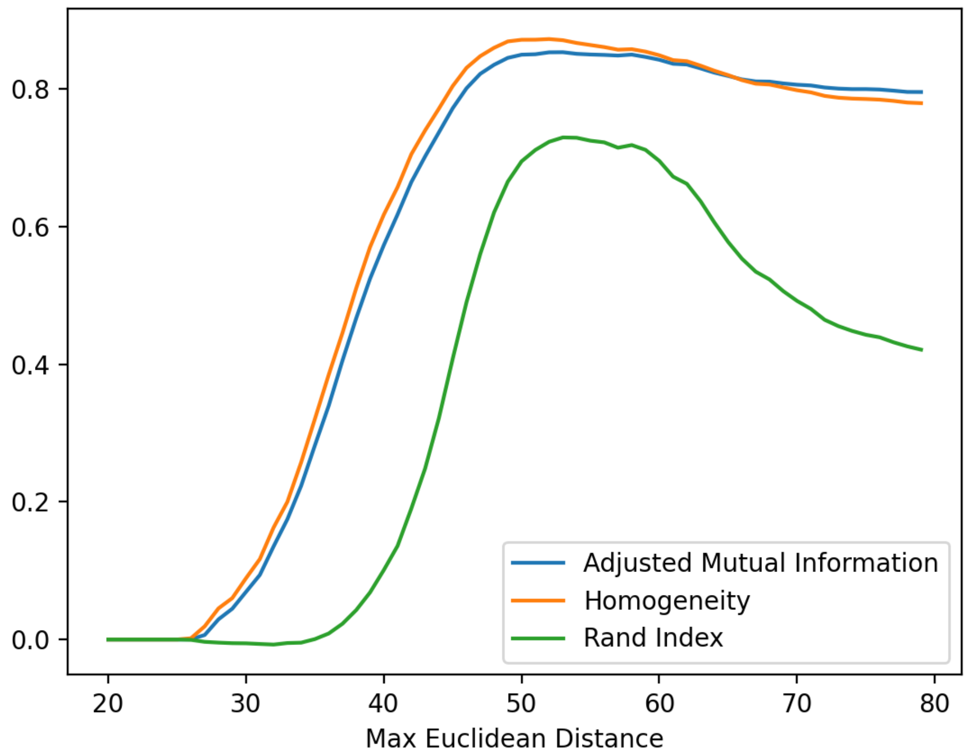
Rows in the FaceDatabase.h5 file now contain a unique numeric identifier, grouping similar faces into clusters. If the “mosaic” option is enabled, an image composed of all of the face tiles in a given cluster is created:

Performance

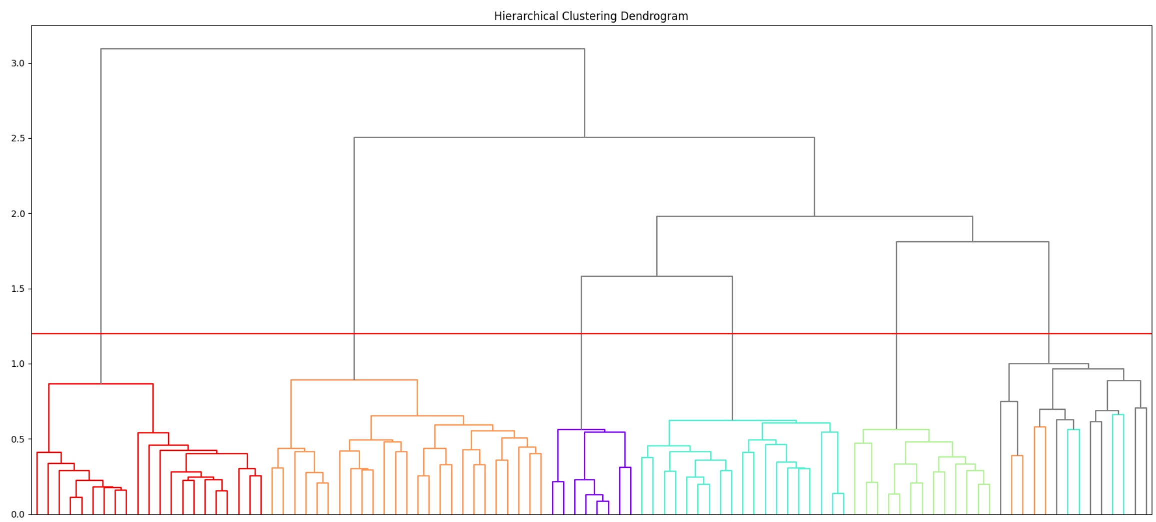
The graph below shows the effect of tuning the “max_distance” parameter on clustering accuracy in the LFW dataset, using the DBSCAN algorithm.

The optimal range seems to be between 50 and 60: if the cutoff is too low, false negatives increase as faces are not sorted into clusters. If the cutoff is too high, the number of false positives increases. Optimal parameters will vary based on your dataset. Further information on the clustering evaluation metrics used above can be found here.

A visual representation of the clustering process is shown in the dendrogram below, generated using Agglomerative Hierarchical Clustering (AHC). Each ”leaf node” (the points where the dendrogram intersects the X axis) represents a face. The Y



axis indicates the euclidean distance (similarity) between facial encodings. If the link between two leaf nodes is very low, the facial encodings are more similar. An interesting property of using AHC to cluster facial images is that the first branch of the dendrogram almost always separates men and women.



The example above uses 100 labeled faces, with colors denoting images of the same individual. We can see groups of faces that are all quite similar to eachother, but quite dissimilar from faces in the other groups. We could probably tell without relying on the colors that there are 5-6 distinct individuals in these 100 images. Indeed, the colors suggest there are 5 main individuals and one "bin" cluster on the far right composed mainly of unlabeled faces (the grey leaf nodes). There are some errant faces from two of the five individuals in this cluster, likely due to poor image quality, pose, or lighting. 94 out of the 100 faces above are labeled. Of these, 86 (91%) were correctly sorted into clusters representing distinct individuals using the red cutoff line.

3. Creating a Network

Having identified individuals across multiple pictures, this function generates a force directed graph based on co-appearance in images. Each individual is a node, and each co-appearance is an edge.

```
face_network.network(photo_dir, scale=10)
```

Parameters

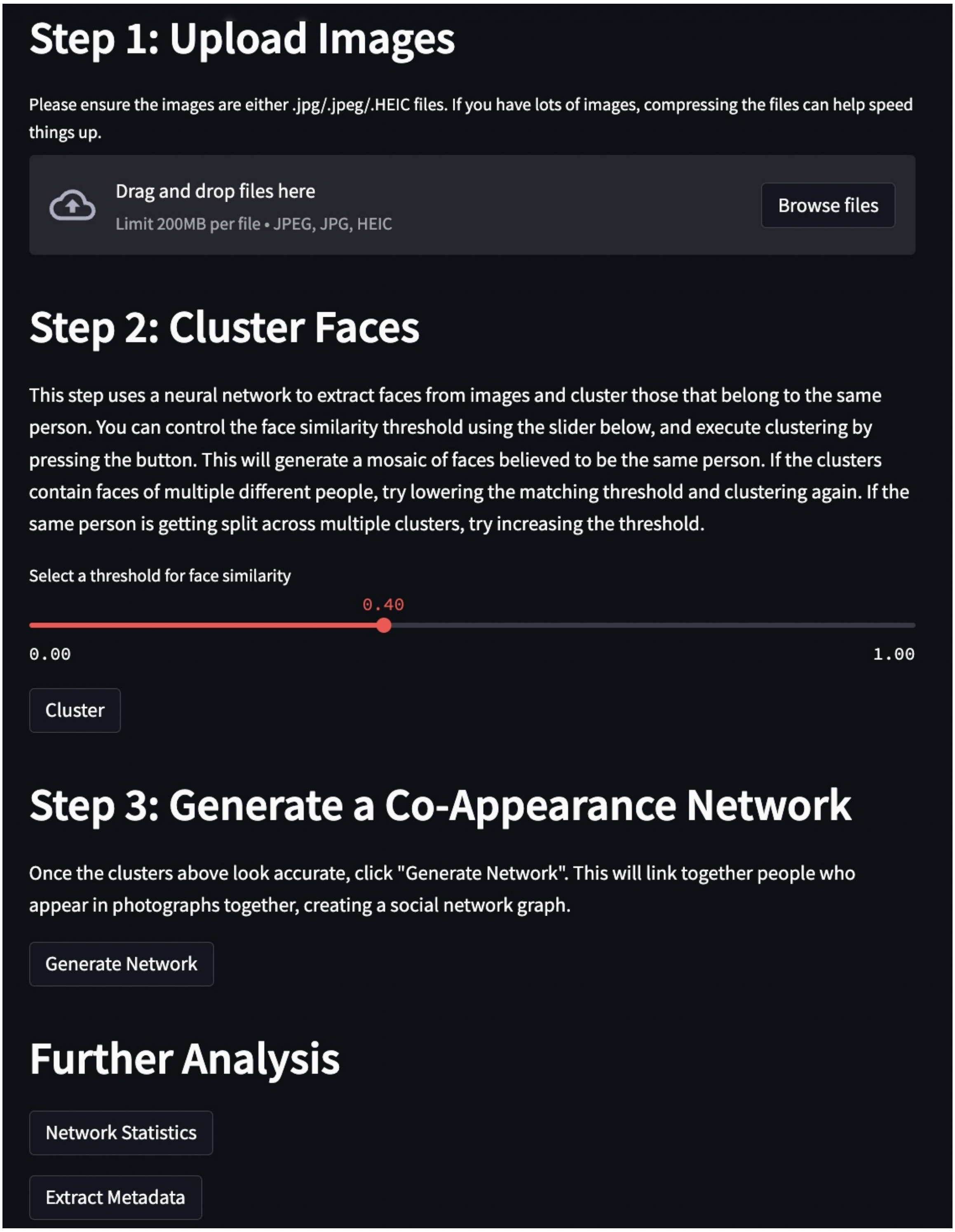
source_dir:	(*str*) The path to the image folder
scale:	(*int, default=10*) Dictates the size of the nodes

Outputs

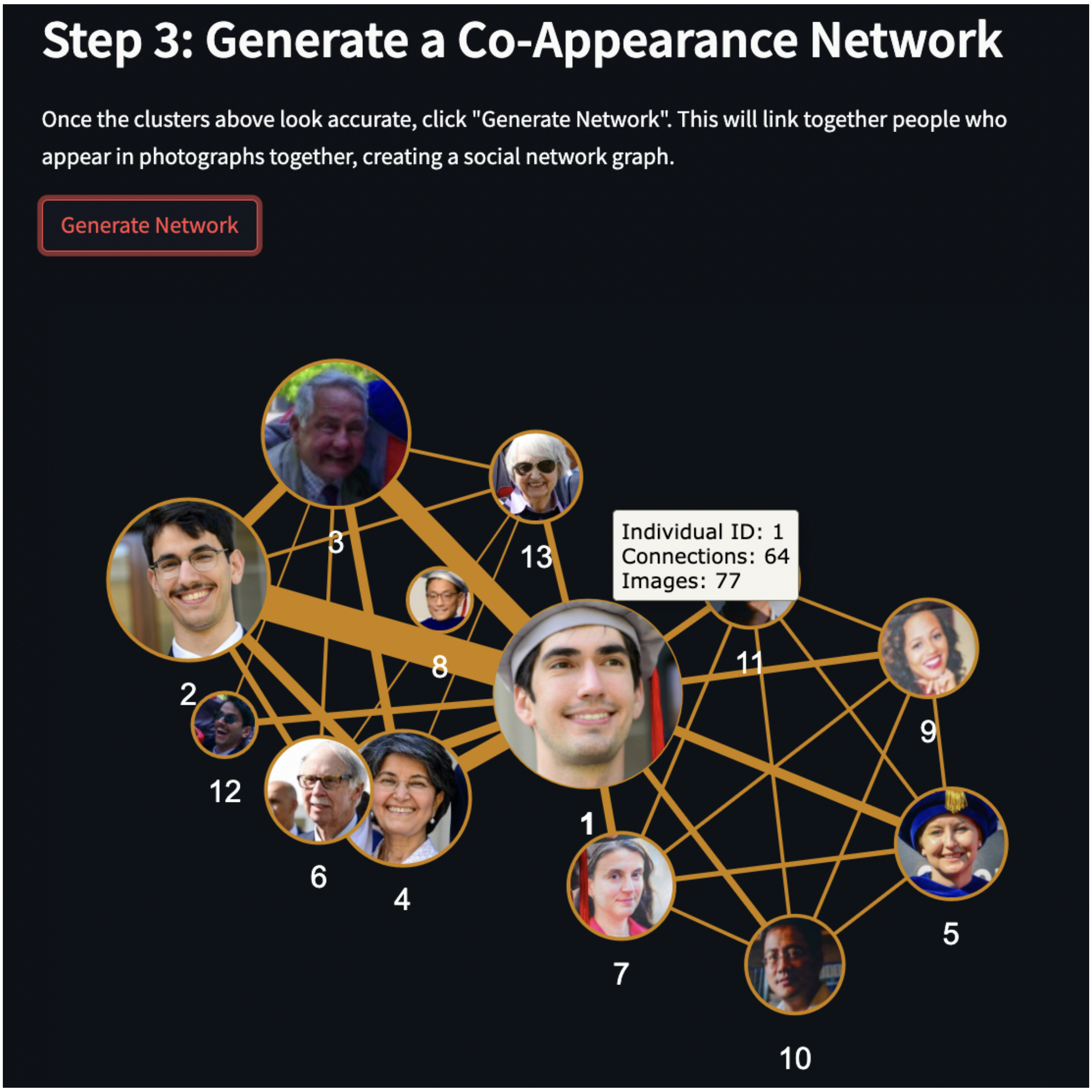
A file called "Image_Network.html" is created in "photo_directory/Face Network/-Data/".The graph can be opened in a web browser and is fully interactive. Hovering over a node will display a tooltip showing the cluster's unique identifier. This corresponds to the filenames of the mosaics generated in the previous step.

The Face-Network Application (with Graphical User Interface)

The Face-Network python package makes the methodology used in Chapter 3 available to a wide range of researchers with basic programming skills. However, in an effort to broaden the accessibility of this methodology even further, I have built an interactive online tool that allows users to use the face-network methodology without coding, by means of a graphical user interface:



In Step 1, users can upload a folder of their own images to the tool. Step 2 allows users to identify individuals across images by clicking the ”cluster” button; the facial similarith threshold can also be toggled by means of a slider in order to fine-tune results to particular datasets. Step 3 generates an interactive graph (via an embedded html file) from the clustering results. Each node has a picture of the individual's face, and a hover-based tooltip lists the number of photographs and co-appearances of an individual. Hovering over an edge between nodes lists the filename of the photograph(s) in which they co-appear. Below is an example of such a graph generated using photographs from my brother's PhD graduation.



Under the ”Further Analysis” heading, network statistics can be generated with the press of a button. These include the total (and LCC) size, number of edges, max degree, clustering coefficient, density, average shortest path, small-world sigma, and assortativity. Finally, metadata contained in the images can be extracted by clicking the "extract metadata" button.

Concluding Remarks

Civil conflicts are complex—their study demands methodological innovation as well as attention to their many facets. Each paper in this thesis presents a new methodological approach to the study of armed groups, focusing on a particular facet of the Kurdish insurgency in Turkey. They each present novel empirical findings on insurgent recruitment, the effects of development policy thereon, and militant group structure. They each present new forward-looking analytical techniques in the study of rebel groups, and make substantive contributions to the literature.

The first chapter analyses patterns in the recruitment of young, predominantly urban recruits using web scraping, fuzzy matching, and other computational techniques. By leveraging an unprecedentedly detailed research design, militants are compared to non-militants to measure regularities in the differences between them with a high degree of precision. This study finds evidence for a range of factors increasing the likelihood of insurgent recruitment, including birth order and family size, peer effects, and conflict-induced migration. This chapter fills two important gaps in the literature on civil conflict. First, it provides rare quantitative empirical evidence on the influence of oft-overlooked social and familial factors in the process of insurgent recruitment. Second, it uses individual-level migratory patterns to contribute to the fraught debate on the role of urbanization and conflict. Despite the high level of precision, an important limitation of this study is that it is correlational, not causal. Nevertheless, the identification of factors linked to insurgent recruitment at the individual level provides a firm basis for further research into their influence beyond the Turkish context.

The second chapter explores economic motivations in Turkey's agrarian Southeast, using remote sensing and spatial econometrics. This provides detailed empirical evidence on one of the main mechanisms linking agricultural income and civil conflict, which speaks directly to the rapidly growing literature on the relationship between rainfall, income, and civil war. This chapter finds that conflict was reduced in areas that benefitted from a state-sponsored agricultural development program. District-level results show that clashes are more frequent following poor harvests, and that irrigation decouples agricultural income from rainfall. By insulating farmers from rainfall shocks and increasing their incomes, irrigation appears to disrupt a key link between climate and conflict. These results align with ethnographic evidence from the area which shows a heightened sense of state legitimacy among beneficiaries of irrigation programs. A necessary qualification is that these effects may be stronger in Turkey than elsewhere given that the PKK was originally a "peasant movement", directly tying many of its grievances to agrarian political economy. Nonetheless, these findings have implications far beyond Turkey, as climate change continues to threaten the livelihoods of rural populations around the globe.

Having examined individual-level motivations for participating in a rebellion, the third chapter examines the PKK as a whole. This chapter develops a new methodology that leverages deep learning to create a social network graph based on co-appearance in photographs which retains many of the broad structural features of the PKK. The primary contribution of this paper is a new methodological approach to the study of insurgent group structure that overcomes the main impediment to this kind of research: "The covert nature of militant groups makes them a difficult subject to study in general, a problem that becomes particularly acute when the objective is to map out the very internal structure that militants go to great lengths to conceal." (Zech and Gabbay, 2016: 231) The source code for this methodology has been published as a Python package, and even an interactive online tool, making it accessible and easy to use by researchers. An important caveat is that the nature of co-appearances is highly contingent on the context in which photographs were taken, and thus care must be taken to interpret the co-appearance network in light of this. Beyond the methodological contribution, the substantive conclusions from this chapter yield novel insights on the trajectory of the conflict between the PKK and the Turkish government. In particular, the densely interconnected nature of this insurgent network suggests that a military approach– even one that targets senior leadership– is unlikely to succeed in dismantling the organization.

There are a number of important limitations to the work presented herein. One of the most fundamental relates the famous adage by Clausewitz that "war is the continuation of politics by other means." People do not join the PKK because the soil is too dry, because their friends told them to, or because their neighbourhood is too crowded. They join, in their own words, due the burning of thousands of their villages, the outlawing of their language and culture, the imprisonment of peaceful political leaders. The factors identified in chapters 1 and 2 may lower the barrier that prevents ordinary people from pursuing politics by extraordinary means, but these should not be conflated with the causes of conflict themselves. Investments in irrigation, education, and social support may blunt the appeal of armed insurrection for some, but as long as grievances are deep and avenues for peaceful political solutions are shallow, the PKK will always have a base of recruits. Treating symptoms is important in both medicine and war, but in neither case should this be confused with a cure.

A final consideration involves the extent to which the lessons learned herein hold elsewhere. Chapter 2 focuses on one particular region within Turkey, and the analysis in chapters 1 and 3 is conducted at the level of individuals. All of this data is interpreted against the backdrop of qualitative work and historical context specific to Turkey. The cost of this level of detail is that the individual findings themselves are contextually contingent and cannot be directly extrapolated to other conflicts. Yet in some sense this is precisely the point of this work. The most widely cited studies of civil war are still those that employ large cross country regressions and are marred by broad and frequently conflicting conclusions. This thesis focuses subnationally on a single conflict, on the individuals who join, the policies aimed at dissuading them from doing so, and their resilience to force when acting in concert. The general framework– leveraging the unprecedented abundance of information being generated, often by militants themselves, to craft a detailed understanding of the causes and conduct of civil conflict– is globally applicable.

I began this thesis with the story of a single rebel, Gülnaz Ekinci, whose life was cut short. My hope is that the methods, analyses, and evidence presented herein contribute towards ensuring that there are fewer stories like hers, not just in Turkey but around the world.

Bibliography

Ahmetbeyzade, C. (2007), 'Negotiating silences in the So-called low-intensity war: The making of the Kurdish Diaspora in İstanbul', *Signs* **33**(1), 159–182. Publisher: The University of Chicago Press.
URL: *https://www.journals.uchicago.edu/doi/abs/10.1086/518315*

Aksit, B. & Akcay, A. A. (1997), 'Sociocultural Aspects of Irrigation Practices in South-eastern Turkey', *International Journal of Water Resources Development* **13**(4), 523–540. Publisher: Routledge _eprint: https://doi.org/10.1080/07900629749601.
URL: *https://doi.org/10.1080/07900629749601*

Anceschi, L., Gervasio, G. & Teti, A., eds (2014), *Informal power in the greater Middle East: hidden geographies*, Routledge studies in Middle Eastern democratization and government, Routledge, Taylor & Francis Group, London ; New York, NY.

Angrist, J. D. & Pischke, J.-S. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press. Google-Books-ID: YSAzEAAAQBAJ.

Arat, Z. F. K. (2007), *Human Rights in Turkey*, University of Pennsylvania Press. Google-Books-ID: iTv7eQ9OQkMC.

Aytekin, M. (2019), 'Radicalisation processes of the Kurdistan Workers Party (PKK): ideology and recruitment tactics', *Journal of Policing, Intelligence and Counter Terrorism* **14**(1), 62–81.
URL: *https://doi.org/10.1080/18335330.2019.1572912*

Barreca, A., Lindo, J. & Waddell, G. R. (2016), 'HEAPING-INDUCED BIAS IN REGRESSION-DISCONTINUITY DESIGNS', *Economic Inquiry* **54**(1), 268–293. Publisher: Western Economic Association International.

URL: *https://econpapers.repec.org/article/blaecinqu/v$_3$a54$_3$ay$_3$a2016$_3$ai$_3$a1$_3$ap$_3$a268–293.htm*

Baser, B. & Ozerdem, A. (2021), 'Conflict Transformation and Asymmetric Conflicts: A Critique of the Failed Turkish-Kurdish Peace Process', *Terrorism and Political Violence* **33**(8), 1775–1796. Publisher: Routledge.
URL: *https://www.tandfonline.com/doi/abs/10.1080/09546553.2019.1657844*

Berry, B. (2006), 'Friends for better or for worse: interracial friendship in the United States as seen through wedding party photos', *Demography* **43**(3), 491–510. Publisher: Demography.
URL: *https://pubmed.ncbi.nlm.nih.gov/17051824/*

Biles, D. (1971), 'Birth order and delinquency', *Australian Psychologist* **6**(3), 189–193.
URL: *https://www.jstor.org/stable/23636636?seq=1*

Bilgen, A. (2018), 'A project of destruction, peace, or techno-science? Untangling the relationship between the Southeastern Anatolia Project (GAP) and the Kurdish question in Turkey', *Middle Eastern Studies* **54**(1), 94–113. Publisher: Routledge _eprint: https://doi.org/10.1080/00263206.2017.1376186.
URL: *https://doi.org/10.1080/00263206.2017.1376186*

Black, S. E., Devereux, P. J. & Salvanes, K. G. (2005), 'The More the Merrier? The Effect of Family Size and Birth Order on Children's Education', *The Quarterly Journal of Economics* **120**(2), 669–700. Publisher: Oxford Academic.
URL: *https://academic.oup.com/qje/article/120/2/669/1933962*

Blake, J. (1989), 'Number of siblings and educational attainment', *Science* **245**(4913), 32–36. Publisher: American Association for the Advancement of Science.
URL: *https://www.science.org/doi/abs/10.1126/science.2740913*

Blattman, C. & Miguel, E. (2010), 'Civil War', *Journal of Economic Literature* **48**(1), 3–57.

BLOMQUIST, W. & SCHLAGER, E. (2005), 'Political Pitfalls of Integrated Watershed Management', *Society & Natural Resources* **18**(2), 101–117. Publisher: Routledge _eprint: https://doi.org/10.1080/08941920590894435.
URL: *https://doi.org/10.1080/08941920590894435*

Bohannon, J. (2009), 'Counterterrorism's new tool: 'Metanetwork' analysis', *Science* **325**(5939), 409–411.
URL: *https://www.science.org/doi/10.1126/science.325*409

Bohlken, A. T. & Sergenti, E. J. (2010), 'Economic growth and ethnic violence: An empirical investigation of Hindu—Muslim riots in India', *Journal of Peace Research* **47**(5), 589–600. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0022343310373032*

Boothby, N., Crawford, J. & Halperin, J. (2006), 'Mozambique child soldier life outcome study: lessons learned in rehabilitation and reintegration efforts.', *Global public health* **1**(1), 87–107.
URL: *https://doi.org/10.1080/17441690500324347*

Bostan, Y. (2016), 'Bahoz Erdal was killed in the third attack - Breaking News'. Publication Title: Sabah.
URL: *https://www.sabah.com.tr/gundem/2016/07/12/bahoz-erdal-ucuncu-saldirida-olduruldu*

Bramoullé, Y., Djebbari, H. & Fortin, B. (2009), 'Identification of peer effects through social networks', *Journal of Econometrics* **150**(1), 41–55. Publisher: North-Holland.

Breland, H. M. (1974), 'Birth order, family configuration, and verbal achievement', *CHILD DEVELOP.* **45**(4), 1011–1019.

Brooks, B., Hogan, B., Ellison, N., Lampe, C. & Vitak, J. (2014), 'Assessing structural correlates to social capital in Facebook ego networks', *Social Networks* **38**(1), 1–15. Publisher: North-Holland.

Buhaug, H. & Urdal, H. (2013), 'An urbanization bomb? Population growth and social disorder in cities', *Global Environmental Change* **23**(1), 1–10.
URL: *https://reader.elsevier.com/reader/sd/pii/S095937801200129X?token=F33DD8022174703 west-1originCreation=20220107180400*

Buolamwini, J. (2018), 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification *', *Proceedings of Machine Learning Research* **81**, 1–15.

Burke, M. A. & Sass, T. R. (2013), 'Classroom peer effects and student achievement', *Journal of Labor Economics* **31**(1), 51–82. Publisher: University of Chicago Press Chicago, IL.
URL: *https://www.journals.uchicago.edu/doi/abs/10.1086/666653*

Calvo, E. & Escolar, M. (2003), 'The Local Voter: A Geographically Weighted Approach to Ecological Inference', *American Journal of Political Science* **47**(1), 189–204. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1540-5907.00013.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-5907.00013*

Chang, L., Pérez-Suárez, A. & González-Mendoza, M. (2019), 'Effective and generalizable graph-based clustering for faces in the wild', *Computational Intelligence and Neuroscience* **2019**. Publisher: Hindawi Limited.

Cheibub, J. A., Gandhi, J. & Vreeland, J. R. (2009), 'Democracy and dictatorship revisited', *Public Choice 2009 143:1* **143**(1), 67–101. ISBN: 1112700994912 Publisher: Springer.
URL: *https://link.springer.com/article/10.1007/s11127-009-9491-2*

Ciccone, A. (2011), 'Economic Shocks and Civil Conflict: A Comment', *American Economic Journal: Applied Economics* **3**(4), 215–227.
URL: *https://www.aeaweb.org/articles?id=10.1257/app.3.4.215*

Collier, P. (2013), *Conflict, political accountability and aid*. Publication Title: Conflict, Political Accountability and Aid.
URL: *https://www.jstor.org/stable/25167719?seq=1metadata*$_{i}$*nfo*$_{t}$*ab*$_{c}$*ontents*

Collier, P. & Hoeffler, A. (1998), 'On economic causes of civil war', *Oxford Economic Papers* **50**(4), 563–573.
URL: *https://doi.org/10.1093/oep/50.4.563*

Collier, P. & Hoeffler, A. (2004), 'Greed and grievance in civil war', *Oxford Economic Papers* **56**(4), 563–595. Publisher: Oxford Academic.
URL: *https://academic.oup.com/oep/article/56/4/563/2361902*

Couttenier, M. & Soubeyran, R. (2014), 'Drought and Civil War in Sub-Saharan Africa', *The Economic Journal* **124**(575), 201–244.
URL: *https://academic.oup.com/ej/article/124/575/201-244/5076982*

Criss, N. B. (1995), 'The nature of PKK terrorism in Turkey', *Studies in Conflict & Terrorism* **18**(1), 17–37. Publisher: Routledge _eprint: https://doi.org/10.1080/10576109508435965.
URL: *https://doi.org/10.1080/10576109508435965*

Dal Bó, E. & Dal Bó, P. (2011), 'Workers, Warriors, and Criminals: Social Conflict in General Equilibrium', *Journal of the European Economic Association* **9**(4), 646–677.
URL: *https://doi.org/10.1111/j.1542-4774.2011.01025.x*

Demuth, S. & Brown, S. L. (2004), 'Family structure, family processes, and adolescent delinquency: The significance of parental absence versus parental gender', *Journal of Research in Crime and Delinquency* **41**(1), 58–81. Publisher: SAGE Publications.
URL: *https://journals.sagepub.com/doi/abs/10.1177/0022427803256236*

Downey, D. B. (1995), 'When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance', *American Sociological Review* **60**(5), 746. Publisher: SAGE Publications.

Downey, D. B. (2001), 'Number of siblings and intellectual development: The resource dilution explanation', *American Psychologist* **56**(6), 497–504. Publisher: American Psychological Association Inc.
URL: */record/2001-17729-003?doi=1*

E. Neumann, R. (2014), 'Bringing the Taliban to the Table: Developing a Framework for Peace in a Country That Has Only Known War'. Issue: 2 Pages: 67–74 Publication Title: Georgetown Journal of International Affairs Volume: 15.
URL: *https://www.jstor.org/stable/43773628*

Edelmann, A., Wolff, T., Montagne, D. & Bail, C. A. (2020), 'Computational Social Science and Sociology', *Annual Review of Sociology* **46**(1), 61–81. _eprint: https://doi.org/10.1146/annurev-soc-121919-054621.
URL: *https://doi.org/10.1146/annurev-soc-121919-054621*

Edgerton, J. (2022), 'Suicide bomber mobilization and kin and peer ties', *Social Networks* **70**, 36–54.
URL: *https://reader.elsevier.com/reader/sd/pii/S0378873321000861*

Elbadawi, E. & Sambanis, N. (2000), 'Why are there so many civil wars in Africa? Understanding and preventing violent conflict', *Journal of African Economies* **9**(3), 244–269.
URL: *https://doi.org/10.1093/jae/9.3.244*

Ellwardt, L., Labianca, G. & Wittek, R. (2012), 'Who are the objects of positive and negative gossip at work? A social network perspective on workplace gossip', *Social Networks* **34**, 193–205.

F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F. & Wilson, R. (2007), 'Methods to account for spatial autocorrelation in the analysis of species distributional data: a review', *Ecography* **30**(5), 609–628. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2007.0906-7590.05171.x.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2007.0906-7590.05171.x*

Fearon, J. D. & Laitin, D. D. (2003), 'Ethnicity, Insurgency, and Civil War', *American Political Science Review* **97**(1), 75–90. Publisher: Cambridge University Press.

URL: *https://www.cambridge.org/core/journals/american-political-science-review/article/ethnicity-insurgency-and-civil-war/B1D5D0E7C782483C5D7E102A61AD6605*

Filkins, D. (2003), 'Kurds Are Finally Heard: Turkey Burned Our Villages - The New York Times'. Publication Title: The New York Times.
URL: *https://www.nytimes.com/2003/10/24/world/kurds-are-finally-heard-turkey-burned-our-villages.html*

Gambetti, Z. & Jongerden, J. (2015), *The Kurdish issue in Turkey: A spatial perspective*, Taylor and Francis Inc. Publication Title: The Kurdish Issue in Turkey: A Spatial Perspective.

GAP (n.d.), 'GAP Nedir? - T.C. GAP Bölge Kalkınma İdaresi Başkanlığı'.
URL: *http://www.gap.gov.tr/gap-nedir-sayfa-1.html*

Garrity, D. L. (1959), 'Family Relationships and Delinquent Behaviour', *Social Forces* **37**(4), 379–380. Publisher: Oxford Academic.
URL: *https://academic.oup.com/sf/article/37/4/379/2227193*

Gatti, N., Baylis, K. & Crost, B. (2021), 'Can Irrigation Infrastructure Mitigate the Effect of Rainfall Shocks on Conflict? Evidence from Indonesia', *American Journal of Agricultural Economics* **103**(1), 211–231. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajae.12092.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/ajae.12092*

Gel, H. P. (n.d.), 'Armanc Kerboran - Hüseyin Akdoğan'.
URL: *http://hpgsehit.com/index.php/ehit-kuenyeleri/2011-ehitlerimiz/item/1497-armanc-kerboran-huseyin-akdogan*

Gizelis, T. I., Pickering, S. & Urdal, H. (2021), 'Conflict on the urban fringe: Urbanization, environmental stress, and urban unrest in Africa', *Political Geography* **86**, 102357. Publisher: Pergamon.

Gleditsch, K. S., Beardsley, K. & Polo, S. M. T. (2010), 'Issues in Data Collection: International Conflict'. ISBN: 9780190846626.
URL: *https://oxfordre.com/internationalstudies/view/10.1093/acrefore/9780190846626.001.0001 9780190846626-e-100*

Golbeck, J. & Klavans, J. L. (2015), *Introduction to Social Media Investigation: A Hands-on Approach*. Publication Title: Introduction to Social Media Investigation: A Hands-on Approach.
URL: *https://books.google.co.uk/books?hl=enlr=id=ZMcSBAAAQBAJoi=fndpg=PP1dq=+Golbe on+approach.+Syngress,+2015.ots=3QVpqgrJArsig=ecBTkiih3U28guKCSqmC8bdkLWkv=onepag*

Goldstone, J. a. (2002), 'Population and Security: How Demographic Change Can Lead to Violent Conflict', *Journal of International Affairs* **56**(1), 3–21. ISBN: 0022-197X.
URL: *https://www.jstor.org/stable/24357881*

Gray, E. (n.d.), 'Drought in eastern Mediterranean worst of past 900 years'.
URL: *https://climate.nasa.gov/news/2408/drought-in-eastern-mediterranean-worst-of-past-900-years*

Grey, J., Brett, R. & Specht, I. (2005), 'Young Soldiers: Why They Choose to Fight', *International Journal* **60**(4), 1181. ISBN: 1-58826-285-5 Publisher: Lynne Rienner Publishers.
URL: *https://books.google.com/books/about/Young$_S$oldiers.html?id* = *mllFkfPKlb4C*

Grover, V. I. (2016), *Water: Global Common and Global Problems*, CRC Press.

Gruen, G. E. (2000), Turkish Waters: Source of Regional Conflict or Catalyst for Peace?, *in* S. Belkin, ed., 'Environmental Challenges', Springer Netherlands, Dordrecht, pp. 565–579.
URL: *https://doi.org/10.1007/978-94-011-4369-1$_4$4*

Gunter, M. M. (2000), 'The continuing Kurdish problem in Turkey after Ocalan's capture', *Third World Quarterly* **21**(5), 849–869.
URL: *https://www.jstor.org/stable/3993622?seq=1metadata$_i$nfo$_t$ab$_c$ontents*

Gurcan, M. (2014), 'Arming civilians as a counterterror strategy: The case of the village guard system in Turkey', *http://dx.doi.org/10.1080/17467586.2014.948026* **8**(1), 1–22. Publisher: Routledge.
URL: *https://www.tandfonline.com/doi/abs/10.1080/17467586.2014.948026*

Güner, S. (1997), 'The Turkish-Syrian war of attrition: The water dispute', *Studies in Conflict & Terrorism* **20**(1), 105–116. Publisher: Routledge _eprint: https://doi.org/10.1080/10576109708436027.
URL: *https://doi.org/10.1080/10576109708436027*

Halverson, J. R., Ruston, S. W. & Trethewey, A. (2013), 'Mediated Martyrs of the Arab Spring: New Media, Civil Religion, and Narrative in Tunisia and Egypt', *Journal of Communication* **63**(2), 312–332. Publisher: Oxford Academic.
URL: *https://academic.oup.com/joc/article/63/2/312/4085995*

Hansen, D. L., Shneiderman, B. & Smith, M. A. (2011), *Analyzing Social Media Networks With NodeXL*, Morgan Kaufmann. Publication Title: Analyzing Social Media Networks with NodeXL.

Harari, M. & Ferrara, E. L. (2018), 'Conflict, Climate, and Cells: A Disaggregated Analysis', *The Review of Economics and Statistics* **100**(4), 594–608.
URL: *https://doi.org/10.1162/rest*$_{a0}$0730

Harris, L. M. (2002), 'Water and Conflict Geographies of the Southeastern Anatolia Project', *Society & Natural Resources* **15**(8), 743–759. Publisher: Routledge _eprint: https://doi.org/10.1080/08941920290069326.
URL: *https://doi.org/10.1080/08941920290069326*

Harris, L. M. (2008), 'Water Rich, Resource Poor: Intersections of Gender, Poverty, and Vulnerability in Newly Irrigated Areas of Southeastern Turkey', *World Development*

36(12), 2643–2662.
URL: *https://www.sciencedirect.com/science/article/pii/S0305750X08001319*

Harris, L. M. (2009), 'States at the Limit: Tracing Contemporary State-Society Relations in the Borderlands of Southeastern Turkey', *European Journal of Turkish Studies. Social Sciences on Contemporary Turkey* (10). Number: 10 Publisher: European Journal of Turkish Studies.
URL: *https://journals.openedition.org/ejts/4122*

Harris, L. M. (2012), 'State as Socionatural Effect: Variable and Emergent Geographies of the State in Southeastern Turkey', *Comparative Studies of South Asia, Africa and the Middle East* **32**(1), 25–39.
URL: *https://read.dukeupress.edu/cssaame/article/32/1/25/59716/State-as-Socionatural-Effect-Variable-and-Emergent*

HARRIS, L. M. (2016), Theorizing gender, ethnic difference, and inequality in relation to water access and politics in southeastern Turkey, *in* 'The Politics of Fresh Water', Routledge. Num Pages: 15.

Hatem, R. & Dohrmann, M. (2013), 'Turkey's Fix for the "Kurdish Problem"', *Middle East Quarterly* . Publisher: Middle East Forum.
URL: *https://www.meforum.org/3667/turkey-kurdish-problem*

Helfstein, S. & Wright, D. (2011), 'Covert or convenient? evolution of terror attack networks', *Journal of Conflict Resolution* **55**(5), 785–813.
URL: *http://jcr.sagepub.com*

Helmers, C. & Patnam, M. (2014), 'Does the rotten child spoil his companion? Spatial peer effects among children in rural India', *Quantitative Economics* **5**(1), 67–121. Publisher: John Wiley & Sons, Ltd.
URL: *https://onlinelibrary.wiley.com/doi/full/10.3982/QE192 https://onlinelibrary.wiley.com/doi/abs/10.3982/QE192 https://onlinelibrary.wiley.com/doi/10.3982/QE192*

Hipel, K. W., Kilgour, D. M. & Kinsara, R. A. (2014), 'Strategic Investigations of Water Conflicts in the Middle East', *Group Decision and Negotiation* **23**(3), 355–376. Publisher: Springer.
URL: *https://econpapers.repec.org/article/sprgrdene/*$v_3a23_3ay_3a2014_3ai_3a3_3ad_3a10.1007_5fs1072$
$012 - 9325 - 3.htm$

Hoeffler, A. & Rohner, D. (2013), Beyond greed and grievance: Feasibility and civil war, *in* 'Conflict, Political Accountability and Aid', pp. 35–62.
URL: *https://www.jstor.org/stable/25167719?seq=1metadata*$_info_tab_contents$

Hogan, D. P. & Blake, J. (1990), 'Family Size and Achievement.', *Social Forces* **69**(1), 305. Publisher: University of California Press.

Holgado Ramos, D. (2016), 'Analyzing Social Networks', *Redes. Revista hispana para el análisis de redes sociales* **27**(2), 141. ISBN: 9781526404107.

Hommes, L., Boelens, R. & Maat, H. (2016), 'Contested hydrosocial territories and disputed water governance: Struggles and competing claims over the Ilisu Dam development in southeastern Turkey', *Geoforum* **71**, 9–20.
URL: *https://www.sciencedirect.com/science/article/pii/S0016718515301020*

Hsiang, S. M., Meng, K. C. & Cane, M. A. (2011), 'Civil conflicts are associated with the global climate', *Nature* **476**(7361), 438–441. Number: 7361 Publisher: Nature Publishing Group.
URL: *https://www.nature.com/articles/nature10311*

Humphries, M. D. & Gurney, K. (2008), 'Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence', *PLOS ONE* **3**(4), e0002051. Publisher: Public Library of Science.
URL: *https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002051*

Jia, Y., Lu, V., Hoberock, J., Garland, M. & Hart, J. C. (2012), 'Edge v. Node Parallelism for Graph Centrality Metrics', *GPU Computing Gems Jade Edition* pp. 15–28. ISBN: 9780123859631 Publisher: Morgan Kaufmann.

Jongerden, J. (2010), 'Dams and Politics in Turkey: Utilizing Water, Developing Conflict', *Middle East Policy* **17**(1), 137–143.
URL: *https://onlinelibrary.wiley.com/doi/10.1111/j.1475-4967.2010.00432.x*

Juárez, N. C., Urdal, H. & Vadlamannati, K. C. (2022), 'The significance of age structure, education, and youth unemployment for explaining subnational variation in violent youth crime in Mexico', *Conflict Management and Peace Science* **39**(1), 49–73.

Kadirbeyoglu, Z. (2017), 'The Impact of Power and Civic Engagement in the Decentralized Management of Natural Resources: The Case of Turkey', *Public Administration and Development* **37**(4), 277–291. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pad.1809.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/pad.1809*

Kadirbeyoğlu, Z. & Özertan, G. (2011), '1 Users ' Perceptions of Water User Associations: Evidence From Three Cases in'.

Kapp, L. (2021), 'Defense Primer: Military Enlisted Personnel Figure 1. Pay Grade, Grade, and Insignia of Enlisted Service Members Resources', *Congressional Research Service (CRS)* .
URL: *https://crsreports.congress.gov*

Kelley, C. P., Mohtadi, S., Cane, M. A., Seager, R. & Kushnir, Y. (2015), 'Climate change in the Fertile Crescent and implications of the recent Syrian drought', *Proceedings of the National Academy of Sciences* **112**(11), 3241–3246.
URL: *https://pnas.org/doi/full/10.1073/pnas.1421533112*

Kemahlıoğlu, (2015), 'Winds of Change? The June 2015 Parliamentary Election in Turkey', *South European Society and Politics* **20**(4), 445–464. Publisher: Routledge.
URL: *https://www.tandfonline.com/doi/abs/10.1080/13608746.2015.1115581*

Klausen, J. (2015), 'Tweeting the Jihad: Social media networks of Western foreign fighters in Syria and Iraq', *Studies in Conflict and Terrorism* **38**(1), 1–22.
URL: *https://www.tandfonline.com/action/journalInformation?journalCode=uter20*

Knoke, D. (2013), '"It Takes a Network": The Rise and Fall of Social Network Analysis in U.S. Army Counterinsurgency Doctrine', *Connections* **33**(1).

Koschade, S. (2007), 'A Social Network Analysis of Jemaah Islamiyah: The Applications to Counterterrorism and Intelligence', *http://dx.doi.org/10.1080/10576100600798418* **29**(6), 559–575. Publisher: Taylor & Francis Group.
URL: *https://www.tandfonline.com/doi/abs/10.1080/10576100600798418*

Kundnani, H. (1999), 'Military action and three deaths after Ocalan's capture'. Publication Title: The Guardian.
URL: *https://www.theguardian.com/world/1999/feb/18/kurds.johnhooper*

Lakomy, M. (2019), 'Recruitment and Incitement to Violence in the Islamic State's Online Propaganda: Comparative Analysis of Dabiq and Rumiyah', *https://doi.org/10.1080/1057610X.2019.1568008* **44**(7), 565–580. Publisher: Routledge.
URL: *https://www.tandfonline.com/doi/abs/10.1080/1057610X.2019.1568008*

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A. & Christakis, N. (2008), 'Tastes, ties, and time: A new social network dataset using Facebook.com', *Social Networks* **30**(4), 330–342. Publisher: North-Holland.

Lordan, O. & Albareda-Sambola, M. (2019), 'Exact calculation of network robustness', *Reliability Engineering and System Safety* **183**(C), 276–280. Publisher: Elsevier.
URL: *https://ideas.repec.org/a/eee/reensy/v183y2019icp276-280.html*

Lyon, A. J. & Uçarer, E. M. (2001), 'Mobilizing ethnic conflict: Kurdish separatism in Germany and PKK', *Ethnic and Racial Studies* **24**(6), 925–948. Publisher: Taylor & Francis.
URL: *https://www.tandfonline.com/doi/abs/10.1080/713766482*

Magouirk, J., Atran, S. & Sageman, M. (2008), 'Connecting Terrorist Networks'.
URL: *https://www.tandfonline.com/action/journalInformation?journalCode=uter20*

Massey, D. S. (1990), 'Social structure, household strategies, and the cumulative causation of migration.', *Population index* **56**(1), 3–26.

Maystadt, J.-F. & Ecker, O. (2014), 'Extreme Weather and Civil War: Does Drought Fuel Conflict in Somalia through Livestock Price Shocks?', *American Journal of Agricultural Economics* **96**(4), 1157–1182. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1093/ajae/aau010.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1093/ajae/aau010*

McCoy, A. B., Wright, A., Kahn, M. G., Shapiro, J. S., Bernstam, E. V. & Sittig, D. F. (2013), 'Matching identifiers in electronic health records: Implications for duplicate records and patient safety', *BMJ Quality and Safety* **22**(3), 219–224. Publisher: BMJ Publishing Group Ltd.
URL: *https://qualitysafety.bmj.com/content/22/3/219 https://qualitysafety.bmj.com/content/22/3/219.abstract*

McKee, T. B., Doesken, N. J. & Kleist, J. (n.d.), 'THE RELATIONSHIP OF DROUGHT FREQUENCY AND DURATION TO TIME SCALES', p. 6.

Medina, R. M. (2014), 'Social Network Analysis: A case study of the Islamist terrorist network', *Security Journal* **27**(1), 97–121. Publisher: Palgrave.
URL: *https://link.springer.com/article/10.1057/sj.2012.21*

Metternich, N. W., Dorff, C., Gallop, M., Weschle, S. & Ward, M. D. (2013), 'Antigovernment Networks in Civil Conflicts: How Network Structures Affect Conflictual Behavior', *American Journal of Political Science* **57**(4), 892–911. Publisher: Blackwell Publishing Ltd.

Miguel, E., Satyanath, S. & Sergenti, E. (2004), 'Economic Shocks and Civil Conflict: An Instrumental Variables Approach', *Journal of Political Economy* **112**(4), 725–753. Publisher: The University of Chicago Press.
URL: *https://www.jstor.org/stable/10.1086/421174*

Morris, J. F. & Deckro, R. F. (2013), 'SNA data difficulties with dark networks', *https://doi.org/10.1080/19434472.2012.731696* **5**(2), 70–93. Publisher: Routledge.
URL: *https://www.tandfonline.com/doi/abs/10.1080/19434472.2012.731696*

Mousseau, D. Y. (2012), 'An inquiry into the linkage among nationalizing policies, democratization, and ethno-nationalist conflict: the Kurdish case in Turkey', *Nationalities Papers* **40**(1), 45–62. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/nationalities-papers/article/abs/an-inquiry-into-the-linkage-among-nationalizing-policies-democratization-and-ethnonationalist-conflict-the-kurdish-case-in-turkey/EC37B1DC4D2324AD7EF729368FC08AD0*

NASA (2013), 'Hyperwall: Ataturk Dam in Turkey from Landsat'.
URL: *https://svs.gsfc.nasa.gov/30218*

Nestor, C. E. (1995), 'Dimensions of Turkey's Kurdish Question and the Potential Impact of the Southeast Anatolian Project (GAP): Part I', *The International Journal of Kurdish Studies* **8**(1/2), 33–II. Num Pages: 47 Place: Brooklyn, United States Publisher: Kurdish Library.
URL: *https://www.proquest.com/docview/216680620/abstract/25E7C1CBDBD34E6APQ/1*

News, A. (2013), 'PKK commemorates Kadir Çelik on the first anniversary of his death'.
URL: *https://anfturkce.com/guncel/pkk-kadir-celik-i-olumunun-birinci-yilinda-andi-23819*

News, A. (2020), 'HSM Command: The blood of our comrade Welat will not be left on the ground'.
URL: *https://firatnews.com/kadin/hsm-komutanligi-welat-yoldasimizin-kani-yerde-kalmayacak-110873*

Ocalan, A. (2013), *Prison Writings - The PKK and the Kurdish Question in the 21st Century (International Initiative Edition)*, TMP Distribution. Google-Books-ID: 2ZYaX07cEUAC.

Ocalan, A. (2015), *Democratic Confederalism*, Lulu Press, Inc. Google-Books-ID: ol5fCAAAQBAJ.

of Interior, T. M. (2022), 'T.C. İçişleri Bakanlığı Terör Arananlar'.
URL: *http://www.terorarananlar.pol.tr/*

of State, U. D. (2022), 'Murat KARAYILAN REWARD Duran KALKAN Up to $5 Million Up to $4 Million Up to $3 Million for Information Cemil BAYIK REWARDS FOR JUSTICE'.
URL: *www.rewardsforjustice.net*

Oktay, Z. (2003), 'WATER DISPUTE AND KURDISH SEPARATISM IN TURKISH-SYRIAN RELATİONS', *The Turkish Yearbook of International Relations* (34), 91–117. Number: 34.
URL: *https://dergipark.org.tr/en/pub/tyir/issue/50002/640952*

Olson, R. (1996), 'The Impact of the Southeast Anatolian Project (GAP) on Kurdish Nationalism in Turkey', *The International Journal of Kurdish Studies* **9**(1/2), 95–102. Num Pages: 8 Place: Brooklyn, United States Publisher: Kurdish Library.
URL: *https://www.proquest.com/docview/216681336/abstract/81F672B479C64EC1PQ/1*

Olson, R. (1997), 'Turkey-Syria Relations since the Gulf War: Kurds and Water', *Middle East Policy* **5**(2), 168–193. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-4967.1997.tb00272.x.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-4967.1997.tb00272.x*

Ozcelik, S. (2006), 'Theories, Practices, and Research in Conflict Resolution and Low-Intensity Conflicts:: The Kurdish Conflict in Turkey', *Journal of Conflict Studies* **26**(2), 133–153. Publisher: The University of New Brunswick.
URL: *https://www.erudit.org/en/journals/jcs/2002-v22-n2-jcs*$_2$6$_2$/$jcs26_2art06$/

Ozdogan, M., Woodcock, C. E., Salvucci, G. D. & Demir, H. (2006), 'Changes in Summer Irrigated Crop Area and Water Use in Southeastern Turkey from 1993 to 2002: Implications for Current and Future Water Resources', *Water Resources*

Management **20**(3), 467–488.
URL: *https://doi.org/10.1007/s11269-006-3087-0*

Patel, R. B. & Burkle, F. M. (2012), 'Rapid urbanization and the growing threat of violence and conflict: A 21 st century crisis', *Prehospital and Disaster Medicine* **27**(2), 194–197. Publisher: Prehosp Disaster Med.
URL: *https://pubmed.ncbi.nlm.nih.gov/22591767/*

Plakoudas, S. (2018), *Insurgency and counter-insurgency in Turkey: The new PKK*. Publication Title: Insurgency and Counter-Insurgency in Turkey: The New PKK.
URL: *https://doi.org/10.1007/978-3-319-75659-2*

Pruitt, L. (2020), 'Rethinking youth bulge theory in policy and scholarship: Incorporating critical gender analysis'. ISSN: 14682346 Issue: 3 Pages: 711–728 Publication Title: International Affairs Volume: 96.
URL: *https://academic-oup-com.eres.qnl.qa/ia/article/96/3/711/5810415*

Pusane, K. (2015), 'Turkey's military victory over the PKK and its failure to end the PKK insurgency', *Middle Eastern Studies* **51**(5), 727–741.
URL: *https://www.tandfonline.com/action/journalInformation?journalCode=fmes20*

Reuters (n.d.), 'Turkish police arrest pro-Kurdish HDP officials'.
URL: *https://www.reuters.com/article/us-turkey-politics-kurds-idUSKBN2BB0AX*

Robins, G., Pattison, P., Kalish, Y. & Lusher, D. (2007), 'An introduction to exponential random graph (p *) models for social networks', *Social Networks* **29**, 173–191.

Rocher, L., Hendrickx, J. M. & de Montjoye, Y. A. (2019), 'Estimating the success of re-identifications in incomplete datasets using generative models', *Nature Communications* **10**(1), 1–9. Publisher: Nature Publishing Group.
URL: *https://www.nature.com/articles/s41467-019-10933-3*

Rodríguez, J. A. (2005), 'The march 11th terrorist network: in weakness lies its strentgh', *Working papers EPP-LEA* **3**.
URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.98.4408*

Saatci, E. & Akpinar, E. (2007), 'Assessing Poverty and Related Factors in Turkey', *Croatian medical journal* **48**(5), 628–635.
URL: *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2205969/*

Sabah, D. (2020), 'PKK founder reportedly killed in Turkish airstrike'.
URL: *https://www.dailysabah.com/politics/war-on-terror/pkk-founder-reportedly-killed-in-turkish-airstrike*

Sageman, M. (2011), *Understanding Terror Networks*, University of Pennsylvania Press.
URL: *https://www.degruyter.com/document/doi/10.9783/9780812206791/html*

Saini, H. S. & Westgate, M. E. (1999), Reproductive Development in Grain Crops during Drought, *in* D. L. Sparks, ed., 'Advances in Agronomy', Vol. 68, Academic Press, pp. 59–96.
URL: *https://www.sciencedirect.com/science/article/pii/S0065211308608433*

Sarsons, H. (2015), 'Rainfall and conflict: A cautionary tale', *Journal of Development Economics* **115**, 62–72.
URL: *https://www.sciencedirect.com/science/article/pii/S030438781400159X*

Scheumann, W. (1998), Conflicts on the Euphrates: An Analysis of Water and Non-water Issues, *in* W. Scheumann & M. Schiffler, eds, 'Water in the Middle East: Potential for Conflicts and Prospects for Cooperation', Springer, Berlin, Heidelberg, pp. 113–135.
URL: *https://doi.org/10.1007/978-3-662-03731-7*$_8$

Schmandt, J. & Kibaroğlu, A. (2016), Sustainability of engineered rivers in arid lands: Euphrates-Tigris and Rio Grande/Bravo, report, TX: LBJ School of Public Affairs. Accepted: 2019-02-15T20:47:57Z.
URL: *http://openaccess.mef.edu.tr/xmlui/handle/20.500.11779/297*

Schroff, F., Kalenichenko, D. & Philbin, J. (2015), FaceNet: A unified embedding for face recognition and clustering, *in* 'Proceedings of the IEEE Computer Society

Conference on Computer Vision and Pattern Recognition', Vol. 07-12-June, pp. 815–823. ISSN: 10636919 _eprint: 1503.03832.

Selby, J., Dahi, O. S., Fröhlich, C. & Hulme, M. (2017), 'Climate change and the Syrian civil war revisited', *Political Geography* **60**, 232–244.
URL: *https://www.sciencedirect.com/science/article/pii/S0962629816301822*

Sidar, R. (2020), 'The rebellious child of the Botan mountains: Reşit Serdar'. Publication Title: ANF News.
URL: *https://anfenglish.com/kurdistan/the-rebellious-child-of-the-botan-mountains-resit-serdar-48045*

Sommers, M. (2011), 'Governance, Security and Culture: Assessing Africa's Youth Bulge', *International Journal of Conflict and Violence* **5**(2), 292–303.
URL: *https://www.ijcv.org/index.php/ijcv/article/view/2874*

Song, L., Chen, M., Gao, F., Cheng, C., Chen, M., Yang, L. & Wang, Y. (2019), 'Elevation Influence on Rainfall and a Parameterization Algorithm in the Beijing Area', *Journal of Meteorological Research* **33**(6), 1143–1156.
URL: *https://doi.org/10.1007/s13351-019-9072-3*

Stansfield, G. & Shareef, M. (2016), 'The Kurdish Question Revisited', p. 256. ISBN: 9781849045919.
URL: *https://books.google.co.uk/books?hl=enlr=id=a3s7DwAAQBAJoi=fndpg=PP1dq=Stansfiel*

Stephens, M. & Poorthuis, A. (2015), 'Follow thy neighbor: Connecting the social and the spatial networks on Twitter', *Computers, Environment and Urban Systems* **53**(53), 87–95. Publisher: Elsevier Ltd.
URL: *https://www.infona.pl//resource/bwmeta1.element.elsevier-35019758-baa5-3ec5-9af2-5dd202f185e7*

Stys, P., Muhindo, S., N'simire, S., Tchumisi, I., Muzuri, P., Balume, B. & Koskinen, J. (2022), 'Trust, quality, and the network collection experience: A tale of two studies on the Democratic Republic of the Congo', *Social Networks* **68**, 237–255. Publisher: North-Holland.

Tapaswi, M., Law, M. & Fidler, S. (2019), Video face clustering with unknown number of clusters, *in* 'Proceedings of the IEEE International Conference on Computer Vision', Vol. 2019-Octob, pp. 5026–5035. ISSN: 15505499 _eprint: 1908.03381.
URL: *https://github.com/makarandtapaswi/BallClustering$_I$CCV*2019

Tastekin, F. (2016), 'Anatomy of a Turkish assassination fable - Al-Monitor: The Pulse of the Middle East'. Publication Title: Al-Monitor.
URL: *https://www.al-monitor.com/originals/2016/07/turkey-syria-pkk-leader-bahoz-erdal-assassination.html*

Tezcür, G. M. (2016), 'Ordinary People, Extraordinary Risks: Participation in an Ethnic Rebellion', *American Political Science Review* **110**(2), 247–264.
URL: *https://www.cambridge.org/core/product/identifier/S0003055416000150/type/journal$_a$rticl*

Tezcür, G. M. (2020), 'A Path out of Patriarchy? Political Agency and Social Identity of Women Fighters', *Perspectives on Politics* **18**(3), 722–739. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/perspectives-on-politics/article/path-out-of-patriarchy-political-agency-and-social-identity-of-women-fighters/F95A18F434CD6254A09D54D2597C05BD*

Trogdon, J. G., Nonnemaker, J. & Pais, J. (2008), 'Peer effects in adolescent overweight', *Journal of Health Economics* **27**(5), 1388–1399. Publisher: North-Holland.

Urdal, H. (2006), 'A clash of generations? Youth bulges and political violence', *International Studies Quarterly* **50**(3), 607–629.
URL: *https://academic.oup.com/isq/article-abstract/50/3/607/1800827*

Urdal, H. (2008), 'Population, resources, and political violence: A subnational study of India, 1956-2002', *Journal of Conflict Resolution* **52**(4), 590–617.
URL: *www.prio.no/cwp/datasets.*

U.S. Department of State (2021), Country Reports on Terrorism 2020, Technical Report April 2019. Publication Title: US Department of Sate.
URL: *https://www.state.gov/reports/country-reports-on-terrorism-2019/PKK*

Uslu, E. (2008), 'Leading PKK Commander Cemil Bayik Crosses into Iran'. Publication Title: Jamestown.
URL: *https://jamestown.org/program/leading-pkk-commander-cemil-bayik-crosses-into-iran/*

Varsamidis, A. (n.d.), 'AN ASSESSMENT OF THE WATER DEVELOPMENT PROJECT (GAP) OF TURKEY: MEETING ITS OBJECTIVES AND EU CRITERIA FOR TURKEY'S ACCESSION', p. 157.

Verma, P. & Zhang, F. (2020), 'The Economics of Telecommunication Services'. ISBN: 978-3-030-33864-0 Place: Cham Publisher: Springer International Publishing.
URL: *http://link.springer.com/10.1007/978-3-030-33865-7*

Verwimp, P., Justino, P. & Brück, T. (2019), 'The microeconomics of violent conflict', *Journal of Development Economics* **141**, 102297.
URL: *https://www.sciencedirect.com/science/article/pii/S0304387818314160*

von Uexkull, N. (2014), 'Sustained drought, vulnerability and civil conflict in Sub-Saharan Africa', *Political Geography* **43**, 16–26.
URL: *https://www.sciencedirect.com/science/article/pii/S0962629814000985*

Warner, J. (2008), 'Contested Hydrohegemony: Hydraulic Control and Security in Turkey', **1**(2), 18.

Warner, J. (2012), 'The struggle over Turkey's Ilısu Dam: domestic and international security linkages', *International Environmental Agreements: Politics, Law and Economics* **12**(3), 231–250.
URL: *https://doi.org/10.1007/s10784-012-9178-x*

White, K. J. C. & Guest, A. M. (2003), 'Community Lost or Transformed? Urbanization and Social Ties', *City & Community* **2**(3), 239–259. Publisher: SAGE PublicationsSage CA: Los Angeles, CA.