Measuring Lethal Counterinsurgency Violence in Amritsar District, India Using a Referral-based Sampling Technique

Romesh Silva* Jeff Klingner Scott Weikart

Abstract
Existing quantitative claims about enforced disappearances in Punjab (India) between 1984 and 1996 are not based on defensible statistical methods. We present preliminary results from a retrospective mortality survey in the rural region of Amritsar District designed to measure lethal counterinsurgency violence. The survey used a hybrid sampling design which combines probability-based random sampling with referral-based sampling. We explore the dynamics of referral-based sampling through a presentation of classical survey estimation methods and social network analysis metrics of the referral chains.

Key Words: conflict mortality, survey methods, respondent-driven sampling, social network analysis, Punjab, India, human rights

1. Introduction

A significant challenge for all empirical research on protracted conflicts is establishing the magnitude, pattern and relative levels of responsibility for "what happened" during a period of mass human rights violations. One of the first questions is always "how many people were killed?" Scientifically-defensible answers assessing the magnitude of the events, and statistical analysis of patterns of pervasive violence, can help to overcome partisan arguments about blame and victimization. Rigorous statistics that quantify violence can de-politicize arguments, help end impunity and begin the process of justice and reconciliation. Recently social scientists have drawn on large-scale empirical data and statistical methods in such fora as international tribunals as well as official truth commissions.[1]

One major challenge for social scientists, when seeking to make scientifically-defensible findings about the magnitude of sexual violence during armed conflict, is the challenge of applying survey research methods to these settings. Both during and after an armed conflict, the population may be unsettled and infrastructure and administrative structures may be damaged. As a result, sampling of population units may be challenging. Further, when seeking to measure mortality retrospectively through survey methods, one needs to employ a form of *indirect sampling* (Lavallée 2007). The need for indirect sampling arises because researchers often have limited or no access to the target population but only access to related populations. In the case of mortality surveys, direct access to the universe of the deceased is not available, so surviving members of the population are sampled and interviewed about the survivorship of their kin.

*All authors are at Benetech, 480 S. California Ave, Suite 201, Palo Alto, CA 94306

[1]For example, Ball (2000a), Ball et al. (2003), and Silva and Ball (2006) used inferential statistical methods to clarify the magnitude and pattern of conflict-related mortality for official truth commissions in Guatemala, Perú and Timor-Leste, respectively. Brunborg (2001), Tabeau et al. (2002), and Ball (2000b), and Ball et al. (2002) contributed expert statistical evidence on killings and refugee flows in cases presented before the International Criminal Tribunal for the Former Yugoslavia.

There are a number of major epistemological challenges to clarifying basic questions about the magnitude and pattern of conflict-related mortality during armed conflict situations. Such challenges stem from issues associated with the measurement of elusive phenomena, the unsettled nature of the population, the difficulties in gaining access to the affected population, and potential security and safety concerns for both enumerators and respondents (Checchi and Roberts 2008).

In this paper, we seek to address some of these challenges through the use of a referral-based sampling technique to retrospectively measure conflict-related mortality. We explore the case of conflict-related mortality, attributable to the State, in rural Amritsar between 1984 and 1996. Our survey uses a hybrid sampling design which combines probability-based random sampling with referral-based sampling. This paper also presents some preliminary analysis of the survey data and its network structure. This preliminary analysis is intended to motivate the extension of classical survey based estimates so that strong assumptions which underpin these estimates can be relaxed.

This paper is structured as follows: in the next section, we provide motivation for this research by presenting a brief overview of the human rights context in Northern India and describing the contours of contentious competing claims (advanced by the government of India and human rights advocates) about the magnitude and pattern of lethal counterinsurgency violence in Punjab. In Section 3, we describe the scientific challenges when attempting to retrospectively estimate direct-conflict mortality during an armed conflict. We note the opportunities and limitations of both triangulation methods and survey-based methods, and argue that these two approaches can be complementary. In Section 4, we describe the design of our hybrid sampling approach to measure direct conflict deaths in the rural region of Amrtisar District, Northern India. We outline how this design provides a means to use the classical Horwitz-Thompson estimator if one makes strong assumptions about the referral process (at the second stage of the sampling). Section 5 presents the resulting estimates using the Horwitz-Thompson estimator, which constitute a lower bound on conflict mortality in Amritsar district during the counterinsurgency. Section 6 discusses the potential for extending this estimate based on analysis of the social network underlying the referral process. In Section 7, we reflect on the challenges involved in adapting existing adaptive/network sampling ideas to retrospectively measure direct conflict-related deaths. We outline four extensions to refine and strengthen our preliminary estimates, before providing some brief concluding remarks in Section 8.

2. Background

Members of the Sikh community comprise 60% of the 24.4 million people in the Northern Indian state of Punjab. Since independence, Sikh community leaders have called for greater autonomy for Punjab, water rights, local control over agricultural production and prices, and official recognition of the Punjabi language, among other concerns. Tension between the Sikh community and the union government in Delhi came to a head in the early 1980s, when the Indian Army deployed approximately 100,000 troops throughout Punjab as part of its counterinsurgency operations, codenamed *Operation Bluestar*. The government's counterinsurgency operations were a response to the increasingly violent acts being carried out by Sikh militants fighting for an independent Sikh nation of Khalistan, including extortion, bombings, murder, and torture (Guha 2006, pp. 557–562).

Human rights groups and media reports have documented that, in the process of carrying out this counterinsurgency, security forces engaged in torture, extrajudicial executions, and enforced disappearances[2] (Kumar et al. 2003). Indian authorities have downplayed reports of abuses in Punjab as "aberrations" that have been addressed according to appropriate procedures.[3] The Government of India's official response to the United Nations Working Group on Enforced and Involuntary Disappearances "denied the allegations that there may be several thousand cases of disappearances in Punjab," and stated, "Scrupulous care had been taken to protect the rights of the individual under due process of law." Furthermore, "wherever there was any suspicion of police excesses, action was taken." (United Nations 1994)

K.P.S. Gill, the director general of police during counterinsurgency operations, described human rights violations as "random excesses" and wrote that "Wherever such excesses were detected, action was inevitably taken. The real question is whether a strategy of State Terrorism was adopted by the police; and the answer is unequivocally in the negative." (Gill 1997) He specifically responded to allegations of enforced disappearances claiming: "Thousands of Sikh youth who had left for foreign countries under fake names and documents were claiming to be missing persons killed by security forces in encounters," adding, "They are missing with the consent of their parents." (Puri 2007).

Human rights groups have rebutted the Indian government's characterizations of abuses as "random excesses" by offering widespread qualitative accounts of human rights violations. These documentation efforts, however, are yet to conclusively clarify whether the government's counterinsurgency efforts were, in fact, a model security operation or an example of the state suspending the rule of law and carrying out mass human rights violations against insurgents and suspected opponents.

Our quantitative data collection and analysis, which includes a referral-based survey in rural Amritsar presented herein, seeks to contribute new empirical evidence and defensible statistical findings to this important human rights debate. We seek to move beyond some of the limitations of existing data and help place this debate on stronger empirical and methodological footing. In particular, we seek to clarify whether the alleged enforced disappearances and extrajudicial executions committed by police and security authorities were "widespread" and/or "systematic." Widespread in that such lethal violence was committed on a large-scale at the population-level, and systematic in the sense that the pattern of these events is consistent with the hypothesis that they are the result of a specific plan or set of official practices.

Our work seeks to build on the efforts of community-based projects, human rights initiatives, and government-sponsored registration processes (Silva et al. 2009). The state of Punjab has a population of over 24 million people and existing qualitative research suggests that there may have been as many as 10–20,000 direct-conflict

[2] *Enforced disappearances* are defined as the arrest, detention, abduction, or any other form of deprivation of liberty by agents of the State or by persons or groups of persons acting with the authorization, support or acquiescence of the State, followed by a refusal to acknowledge the deprivation of liberty or by concealment of the whereabouts of the disappeared person, which place such a person outside the protection of the law (United Nations General Assembly 2006).

[3] At the 50th session of the UN Human Rights Commission in February 1994, Dr. Manmohan Singh, then India's finance minister, downplayed widespread human rights abuses in India as "aberrations" that had occurred in confronting terrorism (Press Trust of India 1994).

deaths attributable to state forces during the 1984–1996 counterinsurgency operations. Hence, as an initial starting point we focus our analysis on the rural areas of Amritsar district to test the scalability of our referral-based sampling technique and explore opportunities to triangulate such referral-based sample data with other existing data sources via multiple systems estimation.

3. Description of Research Problem

3.1 Existing Research on the Problem of Incomplete Data Sources

When evaluating the magnitude and pattern of conflict-related mortality during armed conflicts, social scientists often rely on either available population census data, convenience sample data, retrospective mortality data or a combination of these data sources. By its nature, each convenience source is vulnerable to bias and error. Exclusive reliance on a single convenience sample can produce misleading conclusions about both the actual pattern and the actual magnitude of direct-conflict deaths. As an example, Ball et al. (1999) showed that direct-conflict deaths reported by convenience-sample data systems, such as media reports, were inversely correlated with violent deaths reported in interviews in Guatemala during the most intense years of that conflict. It has also been shown that retrospective mortality surveys of conflict-related deaths that draw on classical survey sampling methods are effective in estimating relatively common conflict-related deaths among the general population. However, they are inefficient and prone to considerable sampling error in estimating deaths that are relatively rare and hidden amongst the general population (Spiegel and Salama 2000). Further evidence of the incomplete and biased nature of data on conflict-related mortality during armed conflict was generated in Timor-Leste by Silva and Ball (2006). They showed that classical survey methods were effective in retrospectively estimating the magnitude and pattern of famine-related deaths common among the general population, but subject to considerable error and uncertainty when used to estimate the highly-targeted phenomena of killings and disappearances (Silva and Ball 2008). Lastly, population census data on communities in affected areas are subject to substantial data quality and coverage defects, which leads to incomplete records (Heuveline 1998). Indirect estimation techniques that draw on population census data are difficult to adjust for the considerable internal and external conflict-related migration that may have occurred. Further, such national-level census data by themselves rarely provide the level of precision needed to examine conflict-related mortality which is concentrated in a particular region of the country or experienced almost exclusively by a small minority or subpopulation.

3.2 The Importance of Developing Quantitative Estimates of Conflict-Related Mortality

When the world considers mass human rights violations, among the first questions is "How many?" Scientifically-defensible answers about the statistical patterns and overall magnitude of massive violence can depoliticize arguments about the past. Rigorous statistics about violence help to overcome partisan arguments about blame and victimhood, instead fostering honest debates about truth, accountability and reconciliation. The moral legitimacy of international human rights and humanitarian work rests on practitioners' claims that they seek to uncover the truth about complex, politicized situations, and to ensure that the human rights and humanitarian response to such situations is guided by an accurate assessment of what happened.

International human rights norms and international humanitarian law have grown over the last 60 years to provide an important set of widely-accepted principles and norms. As Seltzer (2009) notes, these bodies of law provide a detailed framework for understanding and organizing evidence from situations where large-scale human rights violations have taken place. In this vein, a diverse array of professionals from many disciplines have begun to apply quantitative methods to engage basic questions of magnitude and pattern. Such questions include: "How many conflict-related deaths resulted from armed hostilities?", "Does the pattern of conflict-related mortality suggest that these acts were part of a plan or policy by those responsible?" and "Were certain subpopulations of the community more vulnerable, or even targeted by those responsible for the resultant human suffering?" Specialists engaging such questions include statisticians, demographers, epidemiologists, sociologists, political scientists, anthropologists, and economists. The context for this work has been wide-ranging, and includes efforts to organize and improve humanitarian and public-health interventions during armed conflict, the historical clarification of what happened during the conflict, the organizing of evidence for domestic and international prosecutions of those most responsible, and efforts to shape institutional reform of organizations which either directly or indirectly contributed to conflict-related mortality. However, regardless of the varied contexts and forums in which such work is developed, the scientific quantification of mass human rights violations seeks to move beyond partisan rhetoric and informal anecdotes. Such analyses seek to identify errors and biases and quantify the uncertainty. Rigorous quantitative conclusions can be tested according to clear and well-established scientific standards. When quantitative data in the service of human rights fail these standards, the credibility of the human rights enterprise and of the entire victim population can be brought into question. Statistics in the service of human rights must be transparent, reproducible and defensible.

3.3 Previous Survey Based Approaches to Measuring Elusive Phenomena such as Direct Conflict Deaths

Survey implementation is often haphazard and vulnerable to considerable bias during complex humanitarian emergencies, which leads to several methodological limitations. In most crises, the population at risk is rarely stable. Lists of households are non-existent and the residential layout is chaotic, making simple or systematic random sampling difficult. As a result, survey researchers often resort to multi-stage cluster sampling when trying to estimate conflict-related mortality in such situations. Cluster sampling, although easier to implement, is vulnerable to large imprecision and high design effects, which yields large confidence intervals, hampering the interpretation of estimates.

Conflict-related mortality can usually be broken down into direct and indirect conflict-related deaths. Direct deaths refer to those deaths which result from acts of violence such as killings, disappearances and armed hostilities. Whereas indirect deaths refer to those deaths which do not result directly from violent trauma in a conflict, but by the deterioration in health services and food access plus increased risk of disease attributable to the violence. These two types of deaths are rarely distributed in similar ways. Direct-violence deaths tend to be elusive, and targeted at certain particular subpopulations. Whereas indirect-violence deaths, particularly in communities with only basic transportation, health and communications infrastructure, tend to occur at a population-level.

Silva and Ball (2008), in their study of Timor-Leste, have noted the particular issues related to conflict-related mortality in conflict-zones. In particular, they find that classically-designed sample surveys are appropriate for estimating population-based phenomena (such as famine deaths, in the case of Timor-Leste), but are subject to considerable uncertainty for highly targeted and elusive phenomena (such as direct political violence deaths in Timor-Leste).

The development of sample survey techniques to estimate conflict-related mortality in an unsettled environment poses two major challenges:

- Designing a sampling strategy that is able to efficiently document the "hidden" phenomena of direct conflict mortality whereby the selection probability of an individual can be explicitly calculated;
- Developing appropriate survey estimation techniques to estimate conflict-related mortality by sampling surviving relatives of the victims of conflict-related deaths.

Recent advances in the statistical sampling of elusive phenomena seem particularly promising for conflict-related mortality researchers. In particular, adaptive sampling designs and respondent-driven sampling have shown much promise when measuring hidden, vulnerable populations. In contrast to conventional sampling designs, adaptive sampling makes use of values observed in the sample to determine additional units to be sampled.

In adaptive sampling, the sampling statistician specifies

- the initial sampling design (prior to any adaptive sampling)
- the initial sample size
- the description of the neighborhood for a sampling unit
- the condition that triggers or initiates adaptive sampling at a sampled unit

The neighborhood is specified in advance of sampling and is not adaptive. If the y-value (response) of a sampled unit satisfies the condition for adaptive sampling, say, $y \in C$, then the unit's neighborhood is added to the sample. If any other units in that neighborhood satisfy C, then their neighborhoods are also added to the sample. The process continues until a cluster of units is obtained that contains a "boundary" of edge units that do not satisfy C. The final sample consists of n_1, not necessarily distinct, clusters, one for each unit selected in the initial sample. If many of the units satisfy the condition, then the sample could consist of most of the units in the population, and hence be very costly. Thus, the design is most appropriate when the characteristic of interest is highly aggregated or clustered (Thompson 1991).

Adaptive sampling techniques are particularly effective for sampling "hidden populations" for which researchers do not have a reliable sampling frame. Adaptive sampling is a new survey sampling design in which sampling regions, defined as "units," are selected based on values of the variables of interest observed during a sampling survey. In a conventional sampling design, the selection for a sampling unit does not depend on previous observations made during an initial survey; entire sampling units are selected before any physical sampling in the field ever takes place. Therefore, conventional sampling guarantees that the calculated statistics will be unbiased. To use the adaptive sampling technique, however, different estimators must be implemented to guarantee unbiasedness.

The development of sample survey techniques to estimate conflict-related mortality across rural areas of Amritsar District poses two major challenges. First, we must design a sampling strategy that is able to efficiently document the "hidden" phenomena of conflict-related mortality. This requires that the selection probability of an individual death can be explicitly calculated. Then we develop survey techniques to estimate conflict-related mortality from responses of surviving relatives of the victims of conflict-related deaths.

Demographers refer to populations such as those individuals who were disappeared between 1984 and 1996 in Punjab as "hidden populations" or "elusive populations." Hidden populations are defined by two key characteristics. First, they are rare at the "global" population level. Second, there is a lack of comprehensive member registries that could be utilized as a sampling frame to aid estimation of the total magnitude and nature of the hidden population. Sudman et al. (1988) discuss the challenges involved in sampling hidden populations. In particular, they note that the use of classical, probability-based sampling techniques to measure hidden populations are inefficient, expensive and subject to very large standard errors. As a consequence of the weakness of classical sampling methods, we use quasi-adaptive sampling methods.[4]

In this paper, we develop a quasi-adaptive sampling strategy, which combines both probability-based random sampling (at the first sampling stage) and adaptive sampling (at the second stage), to measure the random data omissions of lethal violence in Amritsar. Such a design circumvents the inefficiencies of classical, probability-based random sampling for measuring a hidden population by exploiting the strengths of adaptive sampling (Félix-Medina and Thompson 2004). Furthermore, this design strategy ensures that we can use classical, probability-sampling estimators that are consistent and unbiased, unlike some of the available adaptive sampling estimators. This design utilizes the post-conflict stability of administrative regions in Punjab and the strong social networks of the families of the victims of enforced disappearances.

4. Data and Methods

This survey was conducted jointly by the Benetech Human Rights Data Analysis Group and Ensaaf. It was sponsored by the Echoing Green Foundation and the Sikh Spirit Foundation. The purpose of the survey was to

- assess the magnitude and pattern of lethal violence in rural Amritsar which was part of the State's counterinsurgency in Punjab between 1984 and 1996,

[4]For information on major developments on adaptive/snowball sampling see Goodman (1961), Klovdahl (1989), McKenzie and Mistiaen (2009), Frank and Snijders (1994), Spreen and Zwaagstra (1994), Heckathorn (1997), and Goel and Salganik (2009).

- understand the context and modes of lethal violence used in rural Amritsar during this period, and
- understand the needs and support sought by the families of victims of lethal counterinsurgency violence in rural Amritsar.

4.1 Survey Design

The survey was conducted in July and August 2009. A structured questionnaire, specifically designed and pilot-tested for this study, was used to collect the data.[5] This questionnaire was written in Punjabi and all interviews were conducted in Punjabi. An English version of the survey questionnaire is available at www.hrdag.org/about/india-punjab.shtml. The Punjabi-language version of the survey is available from the authors, upon request. A field survey team of 12 individuals working in teams of two collected the data. Two field supervisors oversaw the survey data collection process.

We designed the survey to study the following reference population: victims of lethal counterinsurgency violence who were killed or disappeared any time between 1984 and 1995 and who were residents of rural Amritsar at the time of their death or disappearance. Given the effects of migration between 1996 and the time of the survey, this sampling plan is expected to result in a downward bias when estimating the magnitude of lethal counterinsurgency violence in rural Amritsar between 1984 and 1996.

We used a referral-based sampling plan. We first randomly selected 190 village clusters proportional to population size (see Section A.1 for details of the sampling frame and sampling procedure). In each sampled village, we then interviewed at least two primary referral points which included local village officials or elders. Primary referral points included:

- Sarpanches: the elected chief administrator of the village
- Bazurgs: groups of village elders who are highly visible in the village
- Chowkidars: village-level government officials who record vital events within the village
- Other village members who are well-known and well-respected within the village. These included granthis (Sikh religious clerics), local shopkeepers, or local school principals.

These primary referral points were asked for referrals to families "who experienced an enforced disappearance and/or extrajudicial execution between 1984 and 1995, and were resident in the sampled village during that time." The survey team then attempted to interview all such families still resident in the village. The survey field team then documented the details of any lethal violence incidents which this family experienced and also asked them for further referrals to other families resident in the sampled village who had experienced acts of lethal violence. All referrals from both primary referral points and families were followed exhaustively.

[5] We conducted six rounds of pilot-testing to refine the questionnaire between August 2007 and May 2009. These rounds of pilot-testing were carried out in three regions of Punjab: Amritsar, Hoshiarpur and Sangrur.

5. Survey Findings

The referral process described in Section 4 is an efficient means of finding rare families who experienced a death during the conflict, but it did not find all of them. Section 5.2.4 discusses some of the reasons why the referral process missed some families, and Section 6 discusses the potential of methods based on social networks to handle this bias. However, classical survey estimates based on the number of victims we did find, based on the assumption that we found all victims, are nonetheless useful as a lower bound on the magnitude of conflict-related mortality.

5.1 Findings from Classical Survey-Based Estimates

By applying the Horwitz-Thompson estimator (Horvitz and Thompson 1952), we estimated that a lower bound of approximately 1,865 on the number of people killed or disappeared in rural parts of Amritsar between 1984 and 1996 by the Indian and Punjab state authorities as part of the State's counter-insurgency campaign against Sikh non-state armed actors. The 95% confidence interval around this point estimate is (1,588, 2,142).[6] This estimate is broadly consistent with convenience sample data sources on lethal state violence in Amritsar District. For example, the Committee for the Coordination on Disappearances in Punjab (CCDP) obtained lists compiled by the police's Central Bureau of Investigation (CBI) about illegal cremations carried out by the Punjab Police in Amritsar District. These CBI lists documented a total of 2,098 bodies which were "illegally cremated" under police orders (after the deceased had been allegedly "disappeared" by the police): 582 fully identified bodies, 278 partially identified bodies, and 1,238 unidentified cremations (Kumar et al. 2003). However, these sources report violent deaths, attributed to the state, that occurred in both rural and urban areas of the district.[7]

In Table 5.1, we present estimates disaggregated by salient demographic and geographic variables. These include estimates of lethal violence attributed to the state disaggregated by whether the victim was Amritdhari,[8] the lethal event occurred in a border village (i.e. a village within 15 km of the Pakistani border), and by the age of the victim.

Table 5.1 shows that, across the rural population of Amritsar district, individuals who were either Amritdhari, reported to be not part of the militancy or between the ages of 20 and 29 were notably more likely to suffer a lethal violence death.

These findings are consistent with some hypotheses advanced by human rights advocates in India (Kumar et al. 2003; Kaur and Dhami 2007; Silva et al. 2009). In particular, our survey estimates support the claims

[6] See Section A.2 for details of the variance estimation method which we used to construct this confidence interval.

[7] Since these convenience samples lack the same precision in geographic reporting as our mortality survey, it is difficult to know exactly what proportion of lethal deaths reported in the *Tribune*, cremation ground records, NHRC documents, etc. occurred in "rural" Amritsar and which occurred in "urban" Amritsar.

[8] An Amritdhari Sikh is someone who has received baptismal vows of the Khalsa initiated by Guru Gobind Singh. An Amritdhari Sikh abides by these vows and follows the *"panj kakari rahit"*, by wearing or displaying five distinctive symbols introduced by the Guru. These include a kesh (long unshorn hair and in case of men, uncut beard), kangha (a comb to keep the hair tidy), kirpan (a sword), kara (a steel bracelet worn about the wrist), and kaccha (a short undergarment).

Disaggregation Variable		Point	Lower	Upper
Characteristic	Value	Estimate	95% CI	95% CI
Border Status	Border Village	738	533	943
	(per 10,000 persons)	9.7	7	12.4
	Non-Border Village	1,127	886	1,369
	(per 10,000 persons)	10.1	7.96	12.3
Amritdhari Status	Amritdhari	1,098	932	1,264
	Non-Amritdhari	757	596	917
Militancy Status	Militant	660	520	799
	Non-Militant	1,198	990	1,406
Age-Group	13–19	299	220	379
	20–29	1,109	911	1,308
	30–39	266	198	335
	40–49	98	63	134
	50+	55	30	80

Table 5.1: Disaggregated Estimates of Killings and Enforced Disappearances that were attributed to the State and occurred in rural Amritsar between 1984-1996

- that more Amritdharis (or baptized Sikhs) were the victims of state-attributed lethal violence than non-Amritdharis,
- that most victims of state-based lethal violence were young males of military age, and
- that the police, although claiming to carry out a focused counterinsurgency against members of the militancy, ended up killing and disappearing a sizable number of individuals who were not known to have any links to the militancy.

However, our survey-based estimates of lethal counterinsurgency violence in border villages compared with non-border villages are inconsistent with the hypotheses of human rights advocates. Human rights advocates have hypothesized that border villages were more vulnerable to direct conflict deaths, theorizing that the police employed the tactic of mass disappearances more frequently in areas close to the Pakistan border. This theory was based on a belief that the Sikh insurgents were strongly supported by the Pakistan intelligence operations from across the border. As Table 5.1 shows, we estimated that border villages and non-border villages experienced similar per-capita levels of lethal counterinsurgency violence. However, our definition of border village is somewhat coarse, as we defined a border village as one which is within 15km of the Pakistan border. Yet it is likely that the notion of border village is more likely to be associated with the police districting boundaries. As such, we plan to carry out post-stratification by this more nuanced definition of "border village."

A key type of policy-relevant question concerns systematic targeting of certain individuals. In order to investigate claims of disproportionate "targeting" of individuals with certain demographic, religious or resident profiles, we would need to assess the relative risk of experiencing lethal violence for these different groups. Such an analysis requires that we have population-level data for the rural parts of Amritsar district disaggregated by these variables. However, such fine-grained population data is not currently available for rural parts of Amritsar district.

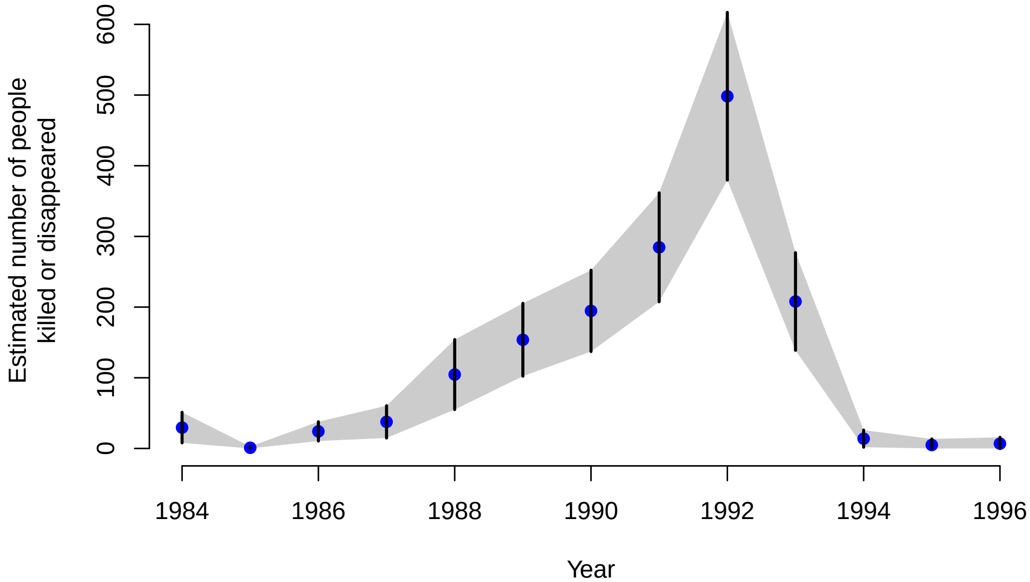


Figure 1: Estimated number of killings and disappearances by year. 16% (79/506) of reported deaths had an unknown date or occurred outside the scope of the survey.

Figure 1 shows our Horwitz-Thompson survey-based estimates disaggregated by year. We estimate that there was a notable and consistent increase in killings and enforced disappearances in rural Amritsar up to 1992, when the phenomena was at its most intense. After 1992, estimated lethal counterinsurgency violence declines rapidly to almost zero by 1994. These estimates are consistent with qualitative reports and existing convenience sample data which suggest that enforced disappearances and extrajudicial executions were overwhelmingly concentrated in the early 1990s when the government intensified its counterinsurgency operations against alleged militants (Silva et al. 2009). The relatively small size of the confidence intervals suggest that despite the phenomena of enforced disappearances and extrajudicial executions being "hidden" or "elusive" at the population level, the hybrid, referral-based sampling scheme was relatively efficient.

5.2 General Limitations of Survey-based Data and Estimates

Almost all data on conflict-related mortality are incomplete. Each data source has its strengths and weaknesses. The challenge for analysts is to understand these strengths and weaknesses and specify appropriate inferential models and methods to account for different types of incompleteness. Convenience sample data on conflict-related mortality is almost always subject to notable types of selection-bias (Gohdes 2010; Ball and Price 2010). Statisticians and demographers have employed capture-recapture and multiple systems estimation methods to try and account for and model such selection biases (Ball et al. 1999). The use of retrospective mortality data to estimate conflict-related mortality involves different, yet related, challenges of incomplete data and reporting bias (Silva and Ball 2008; Checchi and Roberts 2008; Spagat 2010).

The referral-based survey design, which we used in Amritsar, is subject to the following limitations:

- recall bias,
- survivorship bias,
- migration effects, and
- assumptions about the referral process and underlying social networks of the sampled population.

In the following paragraphs, we discuss each of these challenges, in turn.

5.2.1 Recall Bias

All retrospective surveys on mortality are subject to some form of recall bias (Som 1970). In classical demographic surveys, recall lapse by the respondents often manifests itself in the form of rounding error when reporting dates or displacement from one age group to another (Pullum and Stokes 1997). Such classical mortality surveys are characterized by a uniform kin relationship between the respondent and the member of the reference population. For example, in child mortality surveys such as UNICEF's Multiple Indicator Cluster Survey (MICS) program, adult females in sampled households are interviewed about the survivorship of their children (UNICEF 2005). However, in our referral-based sampling design there is notable heterogeneity in the relationship between the respondent and the member of the reference population (i.e. the deceased person). Sometimes, the respondent reports the death of their spouse, other times their husband, brother, son or another member of their kin network. Table 5.2 reports the observed frequencies of kin relations between the respondent and deceased in our referral-based sample. This table show the mix of kin relations reported on by respondents in our referral-based sample. Each bar chart in Table 5.2 shows how often the given date was recalled to a precision of multiple years, one year, one month, or one day. Death dates were recalled with more precision than birth dates. Death date recall was relatively homogeneous across different respondent-victim relationship types, but recall of birth dates was not.

For different kin relations we observed different levels of precision in the reporting of date and age information, as shown in the bar charts in Table 5.2. For example, we note that, on average, victims' parents were able to report on their ages more precisely than their spouses or children. On the other hand, recall of the date of death or disappearance had a similar precision across a variety of relationship types. This variation suggests that recall lapse in this type of survey is somewhat a function of the type of kinship relations between the respondent and the deceased. Given the flexible design of the survey and referral process, this introduces added heterogeneity in the precision of recall and reporting by respondents.

Because the deaths and disappearances occurred 15–25 years before the survey, we included a number of date probes to follow up when the recall of the date was imprecise. We asked respondents if they recalled other information about the date, such as the time relative to the India-Pakistan war, the season, or the month of the Nanakshahi calendar. The Nanakshahi calendar is a twelve-month solar calendar used to track Sikh holy days and is widely used in Punjab. We found that for 19%

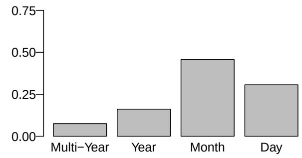
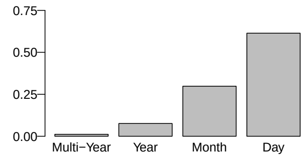
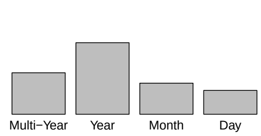
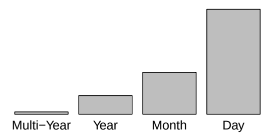
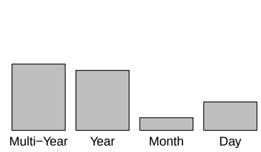
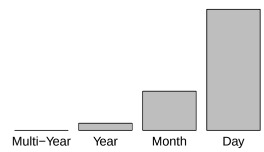
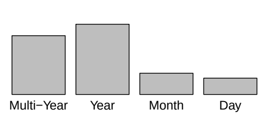
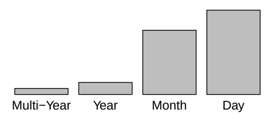
Relationship	Interviews	Recall Precision	
		Birth Date	Death Date
Child	187		
Sibling	163		
Spouse	53		
Other	103		

Table 5.2: The number of interviews and recall precision for respondents of various relationships to the victim they reported.

(94/506) of birth dates and 12% (59/506) of death dates, these probes succeeded in determining the date with more precision than the initial date question.

5.2.2 Survivorship Bias

In retrospective morality surveys, survivorship bias refers to the systematic error which results when deceased individuals have zero probability of being reported in the survey since none of the members of their kinship network survive to the time of the survey. Such survivorship bias has become an increasing challenge in mortality surveys focused on infant/child mortality carried out in high HIV-prevalence countries (Gregson et al. 2009). As systematic bias results from the high correlation between mother and child survivorship, when the mother is HIV-positive. In Timor-Leste, Silva and Ball (2008) noted the survivorship bias challenges of measuring famine-related deaths using a retrospective mortality survey.

Our referral-based survey design is likely to be subject to some survivorship bias. But given the targeted nature of direct conflict deaths during the counterinsurgency in Amritsar, it is unlikely that a notable number of the deceased have no surviving next of kin.

5.2.3 Migration

Our sample was effectively drawn thirteen years after the end of our reference period. During those thirteen intervening years, there has been notable migration into, out of and within Amritsar district.

Rural laborers from outside of Punjab and other parts of the state, seeking higher wages than what is available in Bihar and Rajastan, migrate into rural parts of Amritsar for seasonal agricultural work (Kainth No Date). Also, young individuals and families often migrate (temporarily or permanently) to urban areas for educational or professional opportunities. During the pilot testing for this survey and the actual survey, we also received reports from sarpanches and bazurgs that some families who experienced conflict-related death attributable to the state migrated away from their home village after receiving police harassment as they tried to seek an effective remedy from the State.

Migration of conflict-affected families away form where they were resident may cause disruption to their social networks at the local or village level. As a result, this may affect the reliability and completeness of the referral sampling process.

Nevertheless, the effect of migration on the sampling plan is likely to result in a systematic under-estimation of enforced disappearances and extrajudicial executions. As such, migration effects are likely to lead to more conservative mortality estimates using this referral-based design.

5.2.4 Exhaustive and Error-Free Referrals Process

Our one-stage cluster design assumes that social networks are completely observable and that the survey referral process is exhaustive and error-free. This is most likely an unrealistic assumption, given that social knowledge is likely to be incomplete (due to such factors as recall errors, migration effects in between the end of the counterinsurgency and the survey, etc.) and given that Punjabi society is highly stratified. In addition to potential incompleteness in the referral process from primary referral points, we also need to identify and control for potential biases resultant from when primary referral points may systematically have a tendency to make referrals to particular subpopulations rather than others. This would lead to an over-representation of some subpopulations and an under-representation of others in the estimated total victim population. In short, this would result in a systematic bias in the composition of the reference population.

In the next section, we present preliminary analysis of the dynamics of the referral process and the characteristics of social networks of the families of the disappeared in rural Amritsar. We present a series of descriptive social-network statistics and visualizations which show the heterogeneity of the referral process across different dimensions. These explorations serve to quantitatively characterize the structure of the referral networks. We also study the demographic characteristics (such as age, sex, religion, caste and socioeconomic status) associated with the referral chains within the sampled villages. This analysis of referral dynamics is the basis for determining the coverage rate of the referral-based sampling in villages within which caste and socioeconomic status may influence social knowledge and social relations.

6. Network-based survey estimates

6.1 Review of Social Network Based Estimation Methods

For the reasons outlined in the preceding section, the referral process of this survey (and referral processes in general) cannot enumerate every member of an elusive population. It may nevertheless be possible to estimate the number of victims not interviewed based on the structure of the social networks that drive the referral process. There are two broad classes of survey design and analysis driven by social network analysis: model-based methods and design-based methods.

6.1.1 Model-Based Social Network Estimation

Model-based methods for estimating the size and attributes of populations sampled through referral-driven processes use information about respondents' social network observed during the survey to fit a model of the social network. This network model is then used to estimate the selection probabilities of each observed unit, which can in turn be used in standard unequal-probability estimators such as those by Horvitz and Thompson (1952).

A variety of social network models has been employed for survey estimation, including exponential family random graph models (van Duijn et al. 2009) and latent spatial models (Hoff et al. 2002). Thompson and Frank (2000) describe maximum likelihood estimation of both network and population parameters. This work is extended using Bayesian estimators by Chow and Thompson (2003). Recent work by Handcock and Gile (2010) looks at the problem of missing data (unobserved networks links, realized in our survey as withheld referrals) and possible solutions.

6.1.2 Design-Based Social Network Estimation

In contrast to model-based estimation methods, design-based methods do not attempt to model the structure of the social network which drives the referral process and instead derive sampling probabilities for each respondent based on local network information about that unit, especially its degree, combined with a randomized referral-following procedure which puts a non-zero probability of selection on every member of population (Thompson 1990; Frank and Snijders 1994; Thompson and Seber 1996).

Both model-based and design-based estimators for populations sampled via referral processes make many assumptions about the structure of the network being sampled and the referral process itself. Design-based estimators usually involve stronger assumptions (Kwanisai 2006), but both classes of estimator rely on fairly strong conditions which are not often met in practice, such as:

- Respondents refer you to all other families in the hidden population which they know (perhaps up to a limit imposed by the number of recruitment coupons issued).
- In cases when respondents do not refer everybody they know, they select whom to refer at random.
- In cases when respondents do not refer everybody they know, they accurately report the number of people to whom they *could* have referred. That is, respondents know and accurately report their own degree.
- The network is fully connected.
- The network is reciprocal, i.e. if family A refers to you family B, then family B would also refer you to family A.

Type of Primary Referral Point	Sampled	Consented	Non-Response: Refused Consent	Non-Response: Not Contacted	Referrals Made	In-sample Referrals Followed
Sarpanch	159	140	2	15	283	188
Ex-Sarpanch	64	57	0	6	143	57
Bazurg	216	211	4	1	492	240
Chownkidar	41	38	0	2	65	42

Table 6.1: Summary Statistics of Primary Referral Process

- The probability that two families know each other is determined by attributes of the families (e.g. caste or education level).

Our survey in Punjab did not collect enough information to enable us to calculate network-based survey estimates or to evaluate whether all of the assumptions required by such methods hold for the families of victims of state violence in Punjab. We can nevertheless learn a lot about the referral process and the potential for future network-based estimates by examining the dynamics of the referral process observed in our survey. The next two subsections describe the referral process we observed, including the nature of our primary referral points, the types of referrals they gave, and homophily in the family-to-family referral process that followed.

6.2 Primary Referral Process

Table 6.1 summarizes the primary referral process in the survey. We note that for the two main types of primary referral points, sarpanches and bazurgs, we observe similar in-sample referral rates. This is interesting given that the primary referral points are notably different in terms of their social position and community function: Sarpanches are village officials who are popularly elected through regularly held elections. They effectively are the head of the village, and they chair the village panchayat (or assembly). Many villagers interact with the sarpanch for administrative reasons such as sorting out identity card matters and getting documents authorized. Bazurgs are highly visible sets of village elders. They are often found in central areas of the village, especially around the gurdwara. They are almost always collections of males who have lived for long period times in the village, often for their entire lives. The survey field team sampled a relatively small number of chowkidars as primary sampling points. Chowkidars are village-level government officials who are employed to record births, deaths and marriages within the village. They often carry out other work on the side and can therefore be difficult to locate at short notice.

As discussed, we used a referral-based sample design as a way of indirectly sampling our reference population. Figure 2 shows a referral network from one of the villages in our survey, as an example to illustrate typical referral dynamics.

Tehsil	No. Sampled Villages	Mean No. of Primary Referral Points	Mean Number of Secondary Referral Points
Ajnala	36	2.46	1.85
Amritsar I	22	2.09	2.23
Amritsar II	22	1.76	2.75
Baba Bakala	25	2.46	3.56
Khadur Sahib	17	2.41	4.90
Patti	28	3.07	5.12
Tarn Taran	35	2.54	3.96

Table 6.2: Mean number of Referral points per village cluster for each Tehsil

6.3 Sources of Potential Bias in Link-Tracing Designs

There are a number of sources of potential bias we need to be aware about when we use adaptive link-tracing designs. These include biases resulting from the choice of initial sample, homophily/inbreeding, potential volunteerism of respondents, and oversampling of individuals with large personal network size (Heckathorn 1997). The two sources of potential bias which we will investigate in this paper are: biases resulting from the choice of initial sample and the homophily/inbreeding bias.

6.3.1 Examining biases associated with the choice of initial sample

In many adaptive sampling designs, given that the reference population is elusive and difficult to document, the initial sample is drawn from a convenience sample list. For example, in adaptive link-tracing designs of homeless people in urban areas, researchers often choose their initial sample from administrative lists from local homeless shelters. A critique of this type of design is that the resulting sample is heavily influenced by the initial sample selection, which often is not representative of the reference population. In our referral-based sampling design, we sort to minimize initial sample bias by selecting our initial sample via a 1-stage cluster sample design, whereby village clusters were chosen proportional to the population size and the sample itself was stratified into border village clusters and non-border village clusters. Within each sampled village cluster, we then interviewed primary referral points (namely sarpanches, ex-sarpanches, bazurgs, chowkidars and other village contact points) within these villages as a means to be referred onto families who experienced lethal violence during the counterinsurgency period.

One way to visualize the referral process is as a directed graph or network. When examining such networks, social scientists use measures of centrality to characterize the connectedness of the network being studied. In our referral-based sampling situation, we can use measures such as *closeness* and *outdegree* to examine the relative centrality of the different primary referral points. *Closeness* is the degree an individual is near all other individuals in a sampled village (directly or indirectly). It reflects the ability of the sampling process to access information about lethal violence deaths which occurred in the sampled village through the social network of the referral points. The *degree* is the count of the number of links to other referral points in the social network and is a simple metric of the connectedness of a referral point to families of lethal violence resident in the sampled village.

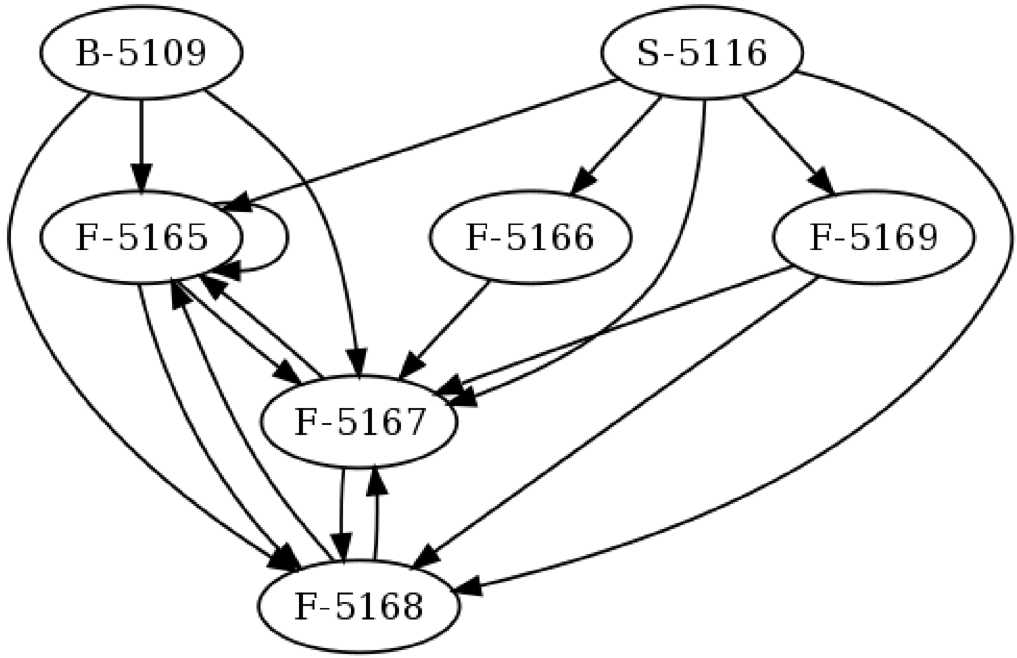


Figure 2: The referral network for the village of Johal Raju Singh, in Tarn-Taran, Punjab. The S- and B- nodes at the top represent the sarpanch and bazurg primary referral points, and the F- nodes represent interviewed families. An arrow from one node to another represents a referral.

As can be seen in Table 6.3, the outdegree of both sarpanches and bazurgs is similar. This indicates that at a coarse and aggregated level these two primary referral points on average have a similar number of links to families of lethal violence resident in the sampled village, whereas chowkidars make less than half as many.

Node Type	Betweenness	Closeness	In degree	Out degree
Bazurg	0.00	0.17	0.00	2.38
Family	1.04	0.01	1.94	0.77
Sarpanch	0.00	0.20	0.00	2.22

Table 6.3: Mean network descriptive statistics by referral network node type.

6.3.2 Homophily bias

Table 6.4 shows that primary referrals by sarpanches were overwhelmingly towards families of the highest Jat caste. Similarly, primary referrals from bazurgs were also overwhelmingly towards high-caste Jat families, regardless of the caste of the bazurg. A key challenge here is to distinguish between caste-based referral bias and potentially disproportionate targeting of people of the Jat caste.

Amongst secondary (or familiy-to-family) referrals, we do observe strong homophily along caste: meaning that family-respondents make referrals to other family-respondents in the same village who are of a similar caste (see Table 6.5).

We also observe strong homophily when the reported victim was a militant, i.e. families of militants have a strong tendency to make referrals to other families of

	Jat	Khatri	Ramgarhia	Mazbi	Dalit	Other	Unknown
Jat	167	0	3	22	0	9	2
Khatri	3	0	0	0	0	0	0
Ramgarhia	0	0	0	0	0	0	0
Mazbi	42	0	1	4	0	5	1
Dalit	3	0	0	0	0	0	0
Other	12	0	0	2	0	3	0
Unknown	190	0	6	32	0	13	7

Table 6.4: Caste homophily of primary referrals

	Jat	Khatri	Ramgarhia	Mazbi	Dalit	Other	Unknown
Jat	265	0	5	14	0	16	2
Khatri	0	0	0	0	0	0	0
Ramgarhia	4	0	1	1	0	0	0
Mazbi	10	0	1	6	0	3	0
Dalit	0	0	0	0	0	0	0
Other	14	0	1	9	0	2	0
Unknown	0	0	0	0	0	0	0

Table 6.5: Caste homophily of secondary referrals

militants who were killed during the counterinsurgency. The converse is not true, however: families of non-militant victims gave about as many referrals to militant as non-militant victims (see Table 6.6).

	Non-militant	Militant
Non-militant	181	62
Militant	57	54

Table 6.6: Militancy homophily of secondary referrals

Yet, we observe that family-respondents are equally likely to make referrals to families with militant victims and non-militant victims irrespective of their political party affiliation.

This exploratory social network suggests that there is notable heterogeneity in the referral process along salient demographic and other characteristics, such as caste and militancy status. Thus assumptions that the referral process is exhaustive and error-free are likely to be unrealistic. Further, the referral process appears to be vulnerable to homophily bias.

7. Future Research Directions

The analysis presented in this paper is preliminary. We have sought to explore the applicability of hybrid sampling designs for the measurement of counterinsurgency-related deaths in rural Amritsar between 1984 and 1996. As an initial means to understand the strengths and weaknesses of such a sampling design used along side classical Horwitz-Thompson estimators, we have drawn on social network analysis to explore the referral chains of the sampling process. As a means to building on this preliminary analysis and formally estimating the magnitude and pattern of lethal violence during the counterinsurgency period in rural Amritsar, we intend to refine and extend this work in a number of directions.

Firstly, we plan to conduct a sensitivity analysis in order to understand the importance of the primary referral points. These respondents played a role similar to the "seed" samples in snowball and RDS surveys, and it is important to learn how much our findings depend on exactly who the survey teams were able to contact as primary referral points. We will consider in turn various subsets of primary referrers: only sarpanches, only bazurgs, only sarpanches and bazurgs, etc, then remove all families only referred to by the removed referrers, all families referred to by those families, and so on, propagating the dropped data through the network of referrals. Then we will recalculate the survey estimates under each of the resulting filtered datasets.

Secondly, as an extension to the preliminary estimates and analysis presented herein, we intend to model missing data patterns in the referral sampling process, to better reflect the imprecision and uncertainty inherent in the referral-based data collection process. Such data imprecision includes the existence of right censored referrals resulting from family out-migration from village, and incomplete referral information (absence of all links, unreliable referral information, and measurement errors). As we develop this part of our research, we intend to draw from recent work by Handcock and Gile (2010) that specifies a maximum likelihood framework while integrating the heterogeneity of precision in the observed network data.

Thirdly, we seek to develop a broader analysis of the survey data. Human rights monitors have argued that the police regularly handed over the bodies of individuals who they "disappeared" to the municipal cremation grounds.[9] The municipal cremation grounds are maintained by the State to dispose of bodies of the deceased whose next of kin are unable to take care of cremation or burial rites. Some existing convenience sample data sources, like the cremation ground records, document cremations events (which followed acts of disappearance) in the municipal cremation grounds, while others document reports of alleged disappearance events. Our referral-based survey data documents the association between alleged disappearances and different body disposal practices, at the population-level. Thus, these survey data are likely to provide new insight into possible connections between these two processes and the chain of custody by perpetrators.

Finally, we seek to integrate our referral-based survey data with the existing data sources on lethal violence and "illegal" cremations. We will do this by deriving multiple systems estimation (MSE) estimates by triangulating our referral-based sample data with existing data sources on lethal violence, collected by the National Human Rights Commission, several non-governmental organizations, press agencies and official government records. This will provide an independent estimate of lethal counterinsurgency violence, which is based on additional data sources and very different methodological assumptions than survey-based estimates. Thus facilitating the estimation of the same parameter by two different methods and complementary sources of data.

[9]Sometimes, the perpetrators allegedly dumped bodies in canals or on the side of roads (Kumar et al. 2003), (Silva et al. 2009).

8. Conclusion

Direct conflict-deaths, at the population-level, are often elusive. As such, when investigating the magnitude and pattern of large-scale direct conflict-related deaths, in certain circumstances we need to draw on and extend the current literature on adaptive sampling designs.

In this paper, we explored the case of direct conflict deaths in rural parts of Amritsar district in the Northern Indian state of Punjab that have been attributed to the State forces. Our focus was the magnitude and pattern of these deaths during the counterinsurgency period of 1984–1996. There exist some diverse convenience sample data sources which document such conflict-related deaths in Amritsar district: municipal cremation ground records, notices of counterinsurgency deaths reported in the local newspapers, (incomplete) registration lists compiled by non-governmental organizations, and an incomplete registration list of families who have applied for compensation from the state. Normally, analysts would use multiple systems estimation to triangulate the data sources and make an inference about the magnitude of deaths never reported to any source, and therefore the total magnitude of the reference population. However, in this particular case, the coverage of the existing data sources is limited and the specificity of the personal identifiers contained in the data sources is imprecise. As a result MSE-based triangulation methods using the existing data sources are insufficient to estimate direct-conflict mortality in Amritsar District. Given the elusive nature of these deaths amongst the broader population, we also noted that classical survey methods were also an insufficient estimation strategy.

We then explored an alternative strategy: a hybrid design which combined elements of classical stratified probability-based random sampling and adaptive sampling designs. Our design, referral-based sampling, exploited the rich social networks of rural society in Amristar District, in general, and those of the families of the disappeared, in particular. We used this hybrid design to overcome some of the traditional sources of bias in earlier snowball-sampling designs. We applied classical Horwitz-Thompson estimators to this design. Our resulting classical estimate was that approximately 1,865 were killed or disappeared in rural parts of Amritsar between 1984 and 1996 by the Indian and Punjab state authorities as part of the state's counter-insurgency campaign against Sikh non-state armed actors. The 95% confidence interval around this point estimate is (1,588, 2,142).

However, this estimate using a classical estimator makes strong assumptions about the completeness of cluster-level enumeration. In particular, it assumed that the referral process was homogeneously reliable and exhaustive. An assumption which is unlikely to be true in the best of times, let alone in a post-conflict rural society in Northern India. We then explored the characteristics of the referral process via standard social network analysis metrics on closeness, betweeness, degree and homophily of referral between primary and secondary referral mechanisms can provide insight into potential biases within the referral networks.

In conclusion, we noted the limitations of classical estimators when applied to hybrid sampling designs. Then, as an extension of our social network analysis of our referral-based sampling design, we outlined some promising future research directions for this work. In particular, we noted opportunities to integrate this work with that of Handcock and Gile (2010) which develops maximum likelihood estimators for network sample data, without assuming error-free and complete network data. We also suggested that our referral-based sampling technique may provide an opportunity to integrate survey estimation with multiple systems estimation methods, when trying to estimate elusive populations such as direct-conflict deaths during the counterinsurgency in Northern India.

A. Appendix

A.1 Village Sampling Procedure

A.1.1 Sampling frame

The first stage of our sample was a probability-proportional-to-size sample of villages in rural Amritsar districted. Our sampling frame was based on a master list of 1,187 villages constructed from the 2001 population census of India, obtained from the Office of the Registrar General & Census Commissioner, India. The list provided the names, locations, number of households, and populations of each village in Amritsar district. The sample was drawn in the rural part of Amritsar only; urban areas (called “wards” in the census) were excluded from the study due to notable sampling frame inaccuracies in urban areas.

A.1.2 Merging villages into clusters before sampling

11% (133/1,187) of the villages had fewer than 50 households. For sampling efficiency, we merged these small villages with neighboring larger villages to form village clusters.

This merging operation resulted in a sampling frame of 1,051 village clusters, 90% (942/1,051) of which comprised a single large village with no nearby small villages, and 10% (109/1,051) of which comprised multiple villages. Most (94) of these multi-village clusters combined a single small village with its neighboring large village. The rest comprised several nearby small villages and a single large neighbor. The largest cluster comprised 7 villages. Exactly one village with more than 50 households was included in each cluster.

A.1.3 Sampling procedure

We used unequal-probability sampling, with the probability of inclusion for each village cluster proportional to the census-reported number of households. We included 5 of the largest villages in Amritsar, where a large number of reported deaths were expected, as certainty samples to reduce sampling variance. These were Sur Singh, Sabhrai, Jhabal Kalan, Sarhali Kalan, and Bundala. None of the certainty samples was a multi-village cluster.

We stratified our sampling on two geographic variables: a binary variable indicating whether the cluster was within 15 km of the border with Pakistan, and a seven-valued variable indicating approximate latitude within Amritsar. Border village clusters occupied latitude zones 1–4, and non-border village clusters occupied

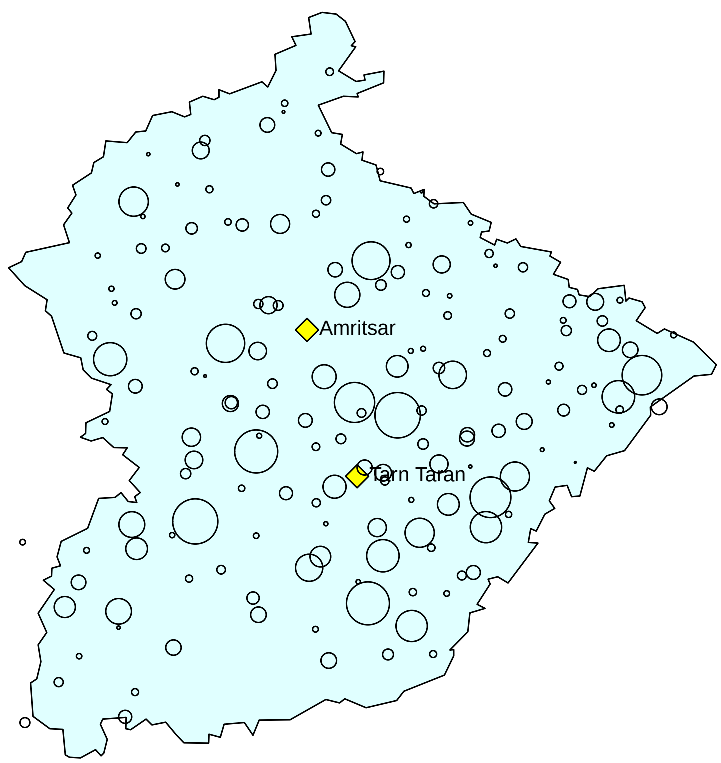


Figure 3: Map of Amritsar district showing locations of sampled village clusters. Each village cluster is shown at the location of GPS coordinates recorded by the interview teams and draw with size proportional to the number of households in the village cluster.

zones 2–7. Locations of each sampled village as reported by the interview teams using GPS receivers are shown in Figure 3.

We selected a sample of 190 village clusters. In order to choose these clusters with probability proportional to their size and stratified geographically as described above, we use a standard randomized-list with random-start systematic sampling procedure, as described by, e.g., Goodman and Kish (1950) and Horvitz and Thompson (1952, p. 678). After randomizing the sampling frame, we sorted it by the two stratification variables, then constructed a cumulative total number of households through each entry of the list. We then stepped through this list in increments of 1668.4 (so there would be 190 steps within the 316,994 total households in the district).

A.2 Estimation

For point estimates, we used the standard Horvitz and Thompson (1952) estimator for the population total under unequal probability sampling:

$$\widehat{Y} = \sum_{i \in \mathcal{S}} \frac{y_i}{\pi_i} \tag{1}$$

where $\mathcal{S}$ is the sample of 190 village clusters, y_i is the number of victims in village cluster i, and π_i is the probability that village cluster i was included in the sample.

A.2.1 Variance estimation

Among many available estimators for the variance of the Horvitz-Thompson estimator, we selected the design unbiased Sen-Yates-Grundy variance estimator (Sen 1953; Yates and Grundy 1953), because it has been found to perform well (Rao and Singh 1973; Cumberland and Royall 1981), and because we meet its requirement of a fixed sample size. This estimator is

$$\widehat{\sigma^2(\widehat{Y})} = \frac{1}{2} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}-i} \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2 \tag{2}$$

The Sen-Yates-Grundy variance estimator requires the joint probabilities of inclusion π_{ij} for each pair of village clusters in our sample. Given our sampling design, these joint probabilities are intractable to compute, so we approximated them as proposed by Stehman and Overton (1994):

$$\pi'_{ij} = \frac{(n-1)\pi_i \pi_j}{n - \frac{1}{2}(\pi_i + \pi_j)} \tag{3}$$

Stehman & Overton showed that this approximation performs as well as the alternative approximation by Hartley and Rao (1962),[10] and that it exhibits especially good performance for populations in which the coefficient of variance of the

[10] There are also alternative variance estimators that do not require pairwise joint probabilities of inclusion (Berger 2004).

auxiliary size variable is less than the coefficient of variance of the response variable, as it was in our survey: cv(number of households per village cluster) = 0.85, cv(number of victims per sampled village cluster) = 1.39.

A.2.2 Confidence intervals

We constructed symmetric 95% confidence intervals for each point estimate of the number of victims as

$$\widehat{Y} \pm 1.973\sqrt{\widehat{\sigma^2(\widehat{Y})}}, \tag{4}$$

where 1.973 is the two-tailed critical value from a t-distribution with $n-1 = 189$ degrees of freedom.

References

Ball, P. (2000a), "The Guatemalan Commission for Historical Clarification: Intersample Analysis," in *Making the Case: Investigating Large Scale Human Rights Violations using Information Systems and Data Analysis*, eds. Ball, P., Spirer, H., and Spirer, L., Washington, DC: American Association for the Advancement of Science, chap. 11.

— (2000b), *Policy or Panic: The Flight of Ethnic Albanians from Kosovo, March–May 1999*, Washington D.C.: American Association for the Advancement of Science.

Ball, P., Asher, J., Sulmont, D., and Manrique, D. (2003), *How many Peruvians have died?*, Washington, DC: American Association for the Advancement of Science.

Ball, P., Betts, W., Scheuren, F., Dudukovich, J., and Asher, J. (2002), *Killings and Refugee Flow in Kosovo, March–June, 1999*, Washington, D.C.: American Association for the Advancement of Science and American Bar Association's Central and Eastern European Law Initiative.

Ball, P., Kobrak, P., and Spirer, H. (1999), *State Violence in Guatemala 1960–1996: a Quantitative Reflection*, Washington, DC: American Association for the Advancement of Science.

Ball, P. and Price, M. (2010), "Unsaid or Unheard: The Limits of Observation and Human Rights Statistics," Benetech HRDAG Working Paper.

Berger, Y. G. (2004), "A Simple Variance Estimator for Unequal Probability Sampling without Replacement," *Journal of Applied Statistics*, 31, 305–315.

Brunborg, H. (2001), "Contribution of Statistical Analysis to the Investigations of the International Criminal Tribunals," *Statistical Journal of the United Nations Economic Commission for Europe*, 18, 227–238.

Checchi, F. and Roberts, L. (2008), "Documenting Mortality in Crises: What Keeps Us from Doing Better?" *PLoS Medicine*, 5, e146.

Chow, M. and Thompson, S. K. (2003), "Estimation with link-tracing sampling designs: A Bayesian approach," *Survey Methodology*, 29, 197–205.

Cumberland, W. G. and Royall, R. M. (1981), "Prediction Models and Unequal Probability Sampling," *Journal of the Royal Statistical Society. Series B (Methodological)*, 43, 353–367.

Frank, O. and Snijders, T. (1994), "Estimating the Size of Hidden Populations Using Snowball Sampling," *Journal of Official Statistics*, 10, 53–67.

Félix-Medina, M. H. and Thompson, S. K. (2004), "Combining Link-Tracing Sampling and Cluster Sampling to Estimate the Size of Hidden Populations," *Journal of Official Statistics*, 20, 19–38.

Gill, K. (1997), "Text of K.P.S. Gill's letter to Prime Minister I.K. Gujral on the death of Ajit Singh Sandhu," South Asia Terrorism Portal, http://www.satp.org/satporgtp/kpsgill/terrorism/97PM.htm (accessed July 25, 2010).

Goel, S. and Salganik, M. J. (2009), "Respondent-driven sampling as Markov chain Monte Carlo," *Statistics in Medicine*, 28, 2202–2229.

Gohdes, A. (2010), "Different Convenience Samples, Different Stories: The Case of Sierra Leone," Benetech Working Paper, available at http://www.hrdag.org/resources/publications/Gohdes_Convenience%20Samples.pdf.

Goodman, L. and Kish, L. (1950), "Control beyond stratification; a technique in probability sampling," *Journal of the American Statistical Association*, 45, 350–372.

Goodman, L. A. (1961), "Snowball Sampling," *The Annals of Mathematical Statistics*, 32, 148–170.

Gregson, S., Hallett, T., Kurwa, F., Garnett, G. P., Dube, S., Chawira, G., Mason, P., and Nyamukapa, C. (2009), "Monitoring progress towards Millennium Development Goal 4 in generalised HIV epidemics: measurement and correction for bias in child mortality statistics," in *Proceedings of the XXVI International Population Conference*, Marrakech, Morocco.

Guha, R. (2006), *India After Gandhi*, New Delhi: Picador.

Handcock, M. S. and Gile, K. J. (2010), "Modeling social networks from sampled data," *Annals of Applied Statistics*, 4, 5–25.

Hartley, H. O. and Rao, J. N. K. (1962), "Sampling with Unequal Probabilities and without Replacement," *The Annals of Mathematical Statistics*, 33, 350–374.

Heckathorn, D. D. (1997), "Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations," *Social Problems*, 44, 174–199.

Heuveline, P. (1998), "Between One and Three Million in Cambodia: Toward the Demographic Reconstruction of a Decade of Cambodian History (1970–1980)," *Population Studies*, 52, 49–65.

Hoff, P. D., Raftery, A. E., and Handcock, M. S. (2002), "Latent Space Approaches to Social Network Analysis," *Journal of the American Statistical Association*, 97, 1090–1098.

Horvitz, D. G. and Thompson, D. J. (1952), "A Generalization of Sampling Without Replacement From a Finite Universe," *Journal of the American Statistical Association*, 47, 663–685.

Kainth, G. S. (No Date), "Migration and Agricultural Development in Punjab," GAD Institute of Development Studies Working Paper, Amritsar, India.

Kaur, J. and Dhami, S. (2007), *Protecting the Killers: A Policy of Impunity in Punjab, India*, New York: Ensaaf and Human Rights Watch.

Klovdahl, A. (1989), "Urban social networks: some methodological problems and possibilities," in *The Small World*, ed. Kochen, M., Norwood, NJ: ABLEX Publishing Corp., pp. 176–210.

Kumar, R. N., Singh, A., Agrwaal, A., and Kaur, J. (2003), *Reduced To Ashes: The Insurgency and Human Rights in Punjab*, Kathmandu. Nepal: South Asia Forum for Human Rights.

Kwanisai, M. (2006), "Estimation in Network Populations," in *Proceedings of the Joint Statistical Meetings*, pp. 3285–3291.

Lavallée, P. (2007), *Indirect Sampling*, Springer series in statistics, New York: Springer, 1st ed.

McKenzie, D. J. and Mistiaen, J. (2009), "Surveying migrant households: a comparison of census-based, snowball and intercept point surveys," *Journal Of The Royal Statistical Society Series A*, 172, 339–360.

Press Trust of India (1994), "India not to submit to terrorism: Manmohan," Tribune Newspaper (Chandigarh), Feb 4.

Pullum, T. and Stokes, S. L. (1997), "Identifying and adjusting for recall error, with application to fertility surveys," in *Survey Measurement and Process Quality*, eds. Lyberg, L. and Paul Biemer, e. a., New York: John Wiley and Sons, pp. 711–732.

Puri, R. (2007), "The Real Fake Encounter," Outlook, http://www.outlookindia.com/article.aspx?234652 (accessed July 25, 2010).

Rao, J. N. K. and Singh, M. P. (1973), "On the Choice of Estimator in Survey Sampling," *Australian & New Zealand Journal of Statistics*, 15, 95–104.

Seltzer, W. (2009), "Conceptual Issues in Quantifying Mortality from Genocide, Crimes against Humanity, and War Crimes," in *Proceedings of the Joint Statistical Meetings, ASA Committee on Scientific Freedom and Human Rights*, American Statistical Association, Alexandria, VA.

Sen, A. R. (1953), "On the estimate of the variance in sampling with varying probabilities," *Journal of the Indian Society of Agricultural Statistics*, 5, 119–127.

Silva, R. and Ball, P. (2006), "The Profile of Human Rights Violations in Timor-Leste, 1974–1999," A Report by the Benetech Human Rights Data Analysis Group to the Commission on Reception, Truth and Reconciliation, available at http://www.hrdag.org/resources/Benetech-Report-to-CAVR.pdf.

— (2008), "The Demography of Conflict-Related Mortality in Timor-Leste (1974–1999): Empirical Quantitative Measurement of Civilian Killings, Disappearances & Famine-Related Deaths," in *Statistical Methods for Human Rights*, eds. Asher, J., Banks, D., and Scheuren, F., New York: Springer, chap. 6, p. 117.

Silva, R., Marwaha, J., and Klingner, J. (2009), "Violent Deaths and Enforced Disappearances During the Counterinsurgency in Punjab, India: A Preliminary Quantitative Analysis," A Joint Report by Benetech's Human Rights Data Analysis Group & Ensaaf, Inc., available at http://www.hrdag.org/about/india-punjab.shtml.

Som, R. K. (1970), *Recall Lapse in Demographic Enquiries*, New York: Asian Publishing House.

Spagat, M. (2010), "Truth and death in Iraq under sanctions," *Significance*, 7, 116–120.

Spiegel, P. B. and Salama, P. (2000), "War and mortality in Kosovo, 1998–99: an epidemiological testimony," *The Lancet*, 355, 2204–2209.

Spreen, M. and Zwaagstra, R. (1994), "Personal Network Sampling, Outdegree Analysis and Multilevel Analysis: Introducing the Network Concept in Studies of Hidden Populations," *International Sociology*, 9, 475–491.

Stehman, S. V. and Overton, W. S. (1994), "Comparison of Variance Estimators of the Horvitz-Thompson Estimator for Randomized Variable Probability Systematic Sampling," *Journal of the American Statistical Association*, 89, 30–43.

Sudman, S., Sirken, M. G., and Cowan, C. D. (1988), "Sampling Rare and Elusive Populations," *Science*, 240, 991–996.

Tabeau, E., Zoltkowski, M., and Bijak, J. (2002), "Population losses in the Siege of Sarajevo, September 10, 1992 to August 10, 1994: expert report prepared for the Galic case (IT-98-29-I)," International Criminal Tribunal for the Former Yugoslavia.

Thompson, S. and Seber, G. (1996), *Adaptive sampling*, New York: Wiley.

Thompson, S. K. (1990), "Adaptive Cluster Sampling," *Journal of the American Statistical Association*, 85, 1050–1059.

— (1991), "Adaptive Cluster Sampling: Designs with Primary and Secondary Units," *Biometrics*, 47, 1103–1115.

Thompson, S. K. and Frank, O. (2000), "Model-based estimation with link-tracing sampling designs," *Survey Methodology*, 26, 87–98.

UNICEF (2005), *Multiple Indicator Cluster Surveys*, New Delhi: United Nations.

United Nations (1994), "Report of the Working Group on Enforced or Involuntary Disappearances, paragraph 222," E/CN.4/1995/36, http://www.unhchr.ch/Huridocda/Huridoca.nsf/TestFrame/5d7027895994334b802566e1005606f9 (accessed 2010-July-5).

United Nations General Assembly (2006), "International Covenant for the Protection of All Persons from Enforced Disappearances," G.A. Res. A/RES/61/177, art. 2.

van Duijn, M. A., Gile, K. J., and Handcock, M. S. (2009), "A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models," *Social Networks*, 31, 52–62.

Yates, F. and Grundy, P. M. (1953), "Selection Without Replacement from Within Strata with Probability Proportional to Size," *Journal of the Royal Statistical Society. Series B (Methodological)*, 15, 253–261.