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 of Timor-Leste.

Romesh Silva and Patrick Ball

Date of Publication: 9 February 2006

Benetech

Romesh Silva and Patrick Ball designed and conducted the statistical analysis and wrote this report. Many Benetech colleagues assisted with this work. Scott Weikart wrote the code to clean datasets, standardize names, de-duplicate datasets and cross-link datasets (sometimes human-assisted, sometimes fully-automated), and he wrote sections of the methodology discussion; Ken Ward wrote the original databases for all three projects; Jana Dudukovic performed hand-matching and hand-cleaning in Palo Alto and in Timor-Leste, and she edited sections of the methodology section; Rafe Kaplan managed the transitions among databases and "flattened" the data to prepare it for analysis; Justin Fisher designed the sampling strategy for the RMS; Sarah Staveteig conducted background research on Indonesian censuses and recoded and organized the displacement data; and Jane Simchuk and Jeff Klingner prepared the report for publication.

In addition, Jana Asher designed the cognitive interviewing component of the RMS questionnaire and provided initial training to the interviewers; Gerry van Klinken had the original idea for the GCD; and Susana Barnes supervised the data coding and data entry teams for all three aspects of the project.

The authors are deeply grateful to our advisers and reviewers who improved our work immeasurably, including Professor David Banks (Duke University), Professor Claes Cassel (Statistics Sweden), Dr Michael Cohen (US Department of Transportation), Dr. Peggy Jennings (Women's Rights International), Richard Öhrvall (Statistics Sweden), Dr. Fritz Scheuren (National Opinion Research Center and American Statistical Association), Professor Herbert F. Spirer (University of Connecticut and Columbia University), William Seltzer (Senior Research Scholar, Fordham University, and American Statistical Association), and Dr. Shana Swiss (Women's Rights International).

This project was funded in part by a subcontract from CAVR to Benetech with funds from the European Union. The authors thank the John D. and Catherine T. MacArthur Foundation, The Oak Foundation, and the Packard Humanities Institute, all of whom provide core support to the Benetech Human Rights Program that was used to complete this project.

The authors retain sole responsibility for the opinions and analysis expressed here, and any errors are theirs alone.

This report should be cited as:
Silva, Romesh and Patrick Ball, "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. 9 February 2006. Available online at http://www.hrdag.org/timor.

Contact Information:
The Benetech Initiative
tel: +1 650-475-5440
fax: +1 650-475-1066
Email: info@benetech.org
Web: http://www.benetech.org

Table of Contents

0. Introduction

In order to achieve the core objectives of its truth-seeking mandate, the Commission developed a number of programs, including both qualitative empirical research and quantitative statistical analysis. This chapter presents the findings that resulted from the Commission's statistical work and discusses the methodological approaches used to reach these findings.

The introductory section, Section 0, provides a brief summary of key findings, an overview of the historical context in which the Commission undertook its demographic and statistical work, and an outline of the Commission's information management decisions.

Section 1 presents the commission's demographic and statistical estimates of the total extent, pattern and trend of, as well as the levels of responsibility for, fatal violations in Timor-Leste between 1974 and 1999. These estimates are derived from statistical and demographic analyses of data collected by the Commission and from external data from official statistical bureaus and human rights NGOs. Although displacement is not necessarily a fatal violation, it is nonetheless closely linked to both conflict-related and famine-related fatalities. Therefore, the analysis of large-scale displacements is included in Section 1.

Section 2 presents an analysis of general patterns of non-fatal violations. The CAVR was unable to find sources containing extensive information about non-fatal violations other than its own testimonies. Consequently we were unable to make rigorous quantitative estimates of the total magnitude of non-fatal violations during 1974-1999. The statistical analysis described in this section therefore focuses on a macro-level view of patterns and trends of non-fatal violations as revealed in the statements given to the Commission.

Section 3 presents a statistical case-study of reported violations experienced by the villagers of Mauxiga in the context of the August 1982 uprising by the resistance movement. This case study is used to highlight the patterns and trends of arbitrary detention, forced displacement and conflict-related deaths which were experienced in the Eastern Region during the consolidation and normalization phase of the Indonesian military's occupation of Timor-Leste.[1]

Section 4 provides a brief summary and conclusion to the chapter.

Section 5 describes the data collection and statistical techniques we used to derive the Commission's statistical findings. It presents methodological background and detailed discussion about the various datasets, data processing methods used, record linkage techniques developed and the analysis and estimation techniques employed.

0.1 Summary of Key Findings

0.1.1 Fatal Violations

We estimate that the minimum-bound for the number of conflict-related deaths during the Commission's reference period, 1974-1999, is 102,800 (+/- 12,000). This estimate is derived from (i) an estimate 18,600 total killings (+/-1000) using multiple systems estimation (MSE)

[1] In this chapter, we define regions in the following way: the Eastern Region comprises Lautem, Viqueque, Baucau and Manatuto. The Central Region comprises Manufahi, Aileu, Dili and Ainaro. The Western Region comprises Ermera, Liquiça, Covalima and Bobonaro.

techniques and (ii) an estimate of 84,200 (+/- 11,000) deaths due to hunger and illness which exceed the total that would be expected if the death rate due to hunger and illness had continued as it was in the pre-invasion peacetime period.

The estimated pattern of fatal violations over time show a high concentration of killings and deaths due to hunger and illness during the initial post-invasion period between 1975 and 1980. The number of deaths attributed by respondents to "hunger or illness" rises to its highest levels during the immediate post-invasion period, 1975-1980. Whereas, 1999 marked the high point for estimated killings 2,634 (+/-626) - which was significantly greater than any other year.

The pattern and trend of deaths due to hunger and illness and killings is positively correlated over time, suggesting that both phenomena have the same underlying cause during the first phase of the conflict. Of the killings and disappearances reported to the Commission's statement-taking process, 57.6% (2,947/5,120) of the perpetrator involvement in fatal violations was attributed to the Indonesian military and Police, and 32.3% (1,654/5,120) to Timorese auxiliaries (such as the militias, civil defense force and local officials who worked under the Indonesian administration).

0.1.2 Displacements

Displacement was widespread: 55.5% of surveyed households reported one or more displacement events, for a total of 2011 reported displacement events between 1974 and 1999.[2]

Most displacements occurred in 1975-1980. The maximum years are 1975 and 1976, with 61,400 (+/- 13,300) and 59,800 (+/- 7,200) displacement events, respectively. The events of 1999 were substantially fewer, with approximately 28,100 (+/- 5,600) events.

Most displacements were local. Of all displacement events, 54.3% are within subdistrict, 15.6% are within district, 17.4% are within region, 9.3%% are within East Timor, and 2.4% are outside of Timor.[3] Many displacements occurred in rapid succession: 22.2% of displacement events lasted one month or less, and 50.1% lasted one year or less. However, other displacements were very long, so that the mean displacement period lasted 46.7 months.[4]

The institution that respondents reported most frequently as the group telling them to move was the Indonesian military (46.4%), followed by FALINTIL (15.0%) and militias (8.8%).[5] Respondents reported that "conflict" motivated 52.3% of their displacements, with "forced by Indonesian military" contributing an additional 16.3%.

0.1.3 Non-Fatal Violations

The temporal pattern of reported non-fatal violations was similar to that for fatal violations: the pattern of massive non-fatal violations during the initial invasion and occupation years, followed by relatively low-level violence during the consolidation and normalization years and then an increase of violence in 1999 is also mirrored in the pattern of fatal violations over time. Whereas the initial violence in the form of non-fatal violations around the time of the Indonesian invasion

[2] According to the 1990 census, there are approximately 4.5 people per household. This figure increases to 4.75 people per household (924,642/194,943) in the 2004 census. The nominal confidence interval is 51.8%-59.2% of households.

[3] The nominal margin of error is +/- 10.4% for within subdistrict, and 4.6% or less for the other estimates. This finding may be limited by the restriction that people in refugee camps in West Timor were not interviewed.

[4] The nominal confidence interval is 41 – 52 months.

[5] The nominal margin of error is +/- 4.2%.

in 1975 was most intense in the Western and Central Regions, after 1976 the focus of non-fatal violations shifted to the Eastern Region.

The observed statistical pattern of reported detentions and tortures suggests that over time (and particularly after 1984) the practice of arbitrary detention became more targeted and was used more regularly in combination with acts of torture. In the early invasion years there are approximately three reported cases of detention for each reported case of torture. After 1985, the two violations appear to be more closely linked, with approximately the same number of reported detentions and reported acts of torture each year.

Overall, the Commission's quantitative findings are consistent with the hypothesis that individuals who were held in detention during the Commission's reference period were subject to increased vulnerability to torture or ill-treatment. Torture and ill-treatment were reported much more frequently among victims who were held in detention during the Commission's reference period: of the torture violations documented by the commission, 83.6% (19,303/11,123) were suffered by victims who had experienced detention during the conflict. The abuses which were most often committed during known periods of detention were torture (38.4%, 4,267/9,094), ill-treatment (33.2%, 27,998/9,094) and threats (21.3%, 634/9,094).

The demographics of victims varied for different violation types. Relative to the overall Timorese population middle-aged males experienced the highest rates of non-fatal violations such as detention, torture and ill-treatment. In contrast, sexually-based violations were almost exclusively targeted against women, with 90.2% (769/853) of reported sexually-based violations being experienced by women.

The Commission's quantitative data suggests a notable difference in the pattern of responsibility for non-fatal violations between 1975 and 1998 relative to non-fatal violations in 1999. In particular, between 1975 and 1998, 51.7% (11,658/22,547) of acts of arbitrary detention are attributed to the Indonesian military acting alone relative to 8.4% (1,897/22,457) of acts of detention which were solely attributed to Timorese auxiliaries or jointly to both the Indonesian occupying force and their Timorese auxiliaries. However, of the acts of arbitrary detention in 1999 documented by the Commission, 75.7% (2,104/2,779) were attributed to either the Timorese auxiliaries acting alone or in collaboration with the Indonesian military and police. Whereas, 19.2% (534/2,779) of documented acts of detention which occurred in 1999 were attributed to the Indonesian military alone.

0.2 Overview of the Commission's Information Management and Data Collection Methods

Most truth commissions base their empirical findings principally on databases derived from the large-scale collection of qualitative testimonies. In this, the CAVR was no different from the commissions in Haiti, South Africa, or Peru. In other countries the truth commissions were able to draw on substantial additional information that had been collected by governmental and non-governmental human rights projects. Additional sources are important in order to "triangulate," or understand the patterns and magnitude of human rights events from perspectives other than the commission's own qualitative material. Without outside corroboration, commissions' work could be dismissed as partisan.

The CAVR did not have massive external sources available, so new sources were created. First, the Commission developed a Human Rights Violations Database (HRVD) from the narrative

testimonies which it collected through direct witness and victim declarations. This was part of the Commission's community socialization process, which sought to address truth-seeking objectives and to promote reconciliation and reception. The Commission used the HRVD narratives for both qualitative and quantitative research.

Second, the Commission developed a retrospective mortality survey (RMS) of 1,396 households that were randomly selected from Timor-Leste's approximately 180,000 households. Each sampled household gave information about their residence pattern and household members and relatives who died during the Commission's mandate period. Mortality surveys of this kind are common among governmental statistical offices to assess health conditions or to adjust censuses. Inter-governmental health authorities and academic demographers and epidemiologists also conduct surveys of this kind. However, no truth commission has ever before conducted a rigorously sampled household survey.

A third dataset collected by the CAVR was the graveyard census database (GCD). Public cemeteries in Timor-Leste were visited, and the name, date of birth and date of death was recorded for every grave for which the information was available. Approximately 327,000 grave records were collected; after duplicate enumerations are removed, there are approximately 319,000 unique graves in the sample, of which about half have complete name and date information. Cemetery records have been used by historical demographers to reconstruct historical patterns of mortality, but no truth commission has ever used data of this kind before as part of the reconstruction of historical memory. In the world of human rights measurement, these are valuable innovations which greatly enrich our understanding of the past.

0.3 Historical Violation Estimates in East Timor and Their Limitations

0.3.1 Historical Estimates of the Conflict-related Death Toll in East Timor (1974-1999)

The scale of conflict-related mortality during Indonesia's occupation of East Timor has been the subject of considerable debate: estimates range from a low of 40,000 to more than 200,000.[7][6] The most informed observers have repeatedly concluded their analyses by recommending that direct evidence be gathered and analyzed. For example, historian Robert Cribb suggests five techniques for measuring total deaths:

- perpetrators' accounts
- counting physical bodies or graves

[6] Estimates based on official Portuguese, Indonesian and Catholic Church data suggest an overall magnitude of approximately 200,000 deaths. See, e.g., Ben Kiernan "The Demography of Genocide in Southeast Asia: The Deathtolls in Cambodia, 1975-79, and East Timor, 1975-80" Critical Asian Studies 35:4 (2003), 585-597, Routledge, and G. Gunn: East Timor and the United Nations: The case for intervention. Lawrenceville, NJ: Red Sea Press. 1997. pp 26-27. On the lower side, see Robert Cribb "How Many Deaths? Problems in the statistics of massacre in Indonesia (1965-1966) and East Timor (1975-1980) in Ingrid Wessel and Georgia Wimhoefer, eds. Violence in Indonesia. Hamburg: Abera-Verl, 2001, Page ?. Waddingham offers a review of estimates derived from "intuitive" and indirect methods, see J. Waddingham "East Timor Death Toll Claims: a Proposal for Listing and Critical Commentary, Submission to the CAVR," 14

- capturing historical memory through interviews
- indirect estimates via census records
- "intuitive" estimates projected from informed observers understanding of local conditions.

Cribb laments that in 1999-2001 at the time that he published a series of papers on the subject, all the available estimates to date, including his own, were limited to the two weakest methods: indirect and intuitive estimates. In his extensive review of the estimates of the conflict-related mortality in Timor-Leste, Waddingham writes "We have to concede, however, that it is not yet possible to produce from available evidence, a quantitatively accurate, generally agreed figure on the death toll in East Timor."[7]

The CAVR was acutely aware of the sensitivity and importance of the estimation of total and disaggregated mortality patterns. Other truth commissions (particularly those in El Salvador, Guatemala and Perú) benefited from the existence of extensive, if still partial, registries of deaths that had been documented before the commission began work. Information of this kind was not available to the CAVR, and so three new datasets were created: a qualitative survey of respondents self-motivated to give testimony to the CAVR; a probability sample of 1396 households from which retrospective mortality histories were taken; and a complete census of all public graveyards in East Timor. These sources fit Cribb's second and third definitions of data sources that could be used to analyze mortality.

While human rights activists sometimes cited the large variation in estimates as evidence of high mortality resulting from the Indonesian occupation of Timor, a possible explanation may be the lack of reliable population and demographic data for the period. The last population census in East Timor before the Indonesian invasion was carried out by the Portuguese colonial administration in 1970. Censuses were conducted by the Indonesian authorities in 1980 and 1990, but the accuracy of these figures is questionable. The Timorese population's suspicion, fear, and general resistance to the government conducting the census combined with the population's frequent movement introduced significant measurement challenges.

Even if the 1980 and 1990 census figures were accurate, their inclusion of only population counts without disaggregating by key demographic variables (such as age and sex) substantially reduces social scientists' ability to apply standard demographic estimation techniques to the official population data. Non-governmental sources of population information are also of limited value due to the severely restricted access independent monitors and humanitarian groups had to East Timor during the conflict. In the context of East Timor, demographer Terence Hull noted that "the variety of estimates in the BPS [Biro Pusat Statistik, Central Statistical Bureau] publications is not an indication of political manipulation of data, but rather the real difficulty of ascertaining mortality levels for small populations when using indirect demographic methods of estimation."[8] Given these limitations, the scientific debate about mortality in East Timor has been unresolved.

[7] J. Waddingham, "East Timor Death Toll, 1975-1999, Submission to the CAVR," 22 July 2003.
[8] Terrence Hull "From Province to Nation: The Demographic Revolution of a People" in James J. Fox, Dionisio Babo Soares, et al "Out of the Ashes: The Destruction and Reconstruction of East Timor" (Crawford House Publishing. Bathurst, Australia: 2000), p38.

Many of the problems facing earlier analysts were resolved by the preliminary publication of the Timor-Leste Census 2004 in March 2005. The CAVR mortality analysis has drawn heavily on the new census data to weight survey results appropriately.

0.3.2 Previous evidence of forced migration and displacement

Large sectors of the population were displaced during the conflict, especially during the early years of the occupation (1975-1980) and in 1999. For example, according to Indonesian official figures, either 268,644 or 318,921 'displaced persons' were held in a total of 15 centers in December 1978.[9] Qualitative reports and the Commission's own research suggests that people were forcibly moved to resettlement camps where a highly restrictive security regime severely limited their opportunities for growing food crops and their access to food sources.[10] As with the estimation of conflict-related fatalities, accurate demographic analysis of displacement in East Timor, especially in the early invasion years is complicated because there were few existing data. This is largely due to the absence of a systematic, country-wide administrative registration system and the limited access to the civilian population available to international humanitarian and human rights organizations during the conflict.

0.3.3 Non-Fatal Violations in East Timor

The extent and depth of information about the human rights situation in East Timor during the Indonesian occupation varied over time. In particular, access to the territory by international human rights monitors was severely restricted by the Indonesian government. Furthermore, the access of international humanitarian agencies (such as the ICRC and Catholic Relief Services) to the territory was restricted to particular areas and particular time periods. when they could work in the territory. International organizations' limited physical access to the territory significantly shaped the international community's knowledge about the human rights situation in Timor-Leste.

1. Analysis of the Total Extent, Pattern, Trend and Levels of Responsibility for Fatal Violations & Displacement in Timor-Leste, 1974-1999[11]

1.1 Background and Overview of Statistical Analysis of Fatal Violations

In the analysis of mortality due to conflict in East Timor, various authors have used varying terms for the manner in which people died. For example, a 2002 essay notes that in Maubisse,

[9] Kohen, Arnold and John Taylor. 1979. "An act of genocide: Indonesia's invasion of East Timor." London: TAPOL. p58.

[10] See CAVR Chapter 7.3: Displacement and Famine, http://www.etan.org/etanpdf/2006/CAVR/07.3_Forced_Displacement_and_Famine.pdf.

[11] The authors gratefully acknowledge comments and suggestions from Professor David Banks (Duke University), Professor Claes Cassel (Statistics Sweden), Dr Michael Cohen (US Department of Transportation), Dr. Peggy Jennings (Women's Rights International), Richard Öhrvall (Statistics Sweden), Dr. Fritz Scheuren (National Opinion Research Center and American Statistical Association), Professor Herbert F. Spirer (University of Connecticut and Columbia University), and Dr Shana Swiss (Women's Rights International).

5,021 of the 9,607 inhabitants alive in 1975 were killed by April 1979.[12] However, the original source for this claim notes the deaths by saying "this village ... has lost" these people, that is, they died in unspecified ways.[13]

The analysis presented here distinguishes between civilian deaths due to violence, called killings; deaths due to hunger and illness, some of which resulted from the conflict; deaths of combatants; and other deaths.[14] Only the patterns and magnitude of killings and deaths due to hunger and illness will be estimated and analyzed.

The distinction between deaths due to hunger and illness and killings is useful for two reasons. First, the person or institution who causes deaths due to deliberate violence has an immediate and obvious responsibility for those deaths, responsibility for deaths due to hunger and illness is more complex. Second, there were substantially fewer killings than deaths by hunger and illness in East Timor between 1975 and 1999. The methods used here to estimate the total number of killings were more precise and underestimate the total less severely than the estimates of the number of deaths due to hunger and illness.

In the sections below, estimates are presented rounded to the nearest hundred in order to signify that the estimates are always approximate and should be interpreted within margins of error. However, specific counts of documented deaths are listed precisely since these numbers are known exactly. The underlying statistical data have been made available so that scholars can continue to analyze the patterns of deaths due to killing and hunger and illness in East Timor during the Indonesian occupation.[15]

Both killings and deaths due to hunger and illness were at higher levels in the immediate post-invasion period, from late 1975 until 1979, than in previous years or at any time until 1999. Both series have peak again in 1999, though killing reaches its highest peak then while hunger and illness have a much smaller peak than in the 1975-1979 period. Our analysis suggests that during the period 1975-1999, approximately 18,600 non-combatants were killed. Furthermore, we estimate that more than 84,200 people died in excess of the peacetime baseline rates of death due to hunger and illness.

The two series – killings and deaths due to hunger and illness – follow the same pattern: the correlation coefficient between the annual estimates (described below) of the number of killings and deaths due to hunger and illness is 0.81, a very high level for most social science findings.[16] Analysis of patterns of displacement suggest that displacement was at its highest levels in the period 1975-1979. The high correlation between estimated numbers of killings, deaths due to hunger and illness, and displacement suggests that they are responding to similar underlying

---

[12] John G. Taylor, "'Encirclement and Annihilation': The Indonesian Occupation of East Timor," chapter 8 in The Specter of Genocide: Mass Murder in Historical Perspective, ed by Robert Gelletely and Ben Kiernan. Cambridge, UK: Cambridge UP

[13] Report on East Timor," East Timorese Church document, 12 July 1979, cited in C. Gilbert and J. Waddingham, "East Timor – How many people are missing?" A report by the Timor Information Service to the Australian Senate Standing committee on Foreign Affairs and Defence, 28 March 1982.

[14] The "other" category included accidents and maternal mortality, among other causes. These deaths were not identified specifically in the questionnaire.

[15] See Timor_Leste data (http://www.hrdag.org/resources/timor-leste_data.shtml) for copies of the statistical data; note that there is no personally identifiable information about witnesses, victims, or perpetrators in this data.

[16] The correlation relates the MSE estimated annual number of killings to the RMS estimated number of deaths due to hunger and illness.

conditions. That is, the three phenomena are likely to have a common cause. The pattern of rapid increase in killings, deaths due to hunger and illness, and displacement at the beginning of the Indonesian occupation is consistent with the claim that the occupation caused the increased mortality.

1.2 Objectives of Analysis

The analysis begins with an overview of the data and methods relevant to the analysis of fatal violations and displacement. In Section 1.4, the estimates of the number of killings and deaths due to hunger and illness are presented. For each manner of death, RMS and MSE estimates are presented and compared. The displacement section reviews the estimated total displacement events and the number of displaced households over time and space. In Section 1.5, descriptive analysis of the deaths reported to the CAVR in the HRVD is presented. Analysis considers the patterns over time, space, collective deaths, demography and political affiliation of victims, institutional responsibility, and the relationship between detention and conflict deaths.

1.3 Overview of Data and Methods

The estimates of the patterns and magnitude of mortality are based on three original data sources collected by the Commission (referred to here as the CAVR), including:

- A collection of qualitative testimonies by respondents self-motivated to give their stories to the CAVR (denoted as the Human Rights Violations Database, HRVD). In this project, 7,668 respondents recounted narratives about violations they suffered or witnessed during the 1974-1999 period. The respondents selected themselves to give reports to the CAVR. Therefore, the results of this project may not represent the entire universe of all people who suffered human rights violations.
- A probability sample of 1,396 households from which retrospective mortality histories were recorded (denoted RMS). In each household, two adult respondents were chosen at random. For male respondents, the respondent's parents and siblings were enumerated, including whether they are living or dead. If they were dead, the date, place, and manner of death were recorded. For female respondents, her children were enumerated in a similar fashion. In 60% of the households, only respondents of one sex were available at the time of the interview, and so one respondent gave information about parents, siblings, and children.
- A complete census of all public graveyards in East Timor, documenting more than 319,000 graves (denoted GCD).

Each data source documents only a small fraction of the total deaths in East Timor, 1975-1999. Even in the absence of conflict, not all of the dead are buried in public graveyards: some people are buried in remote locations or in private family graveyards. When mortality conditions are especially severe, relatively fewer people are buried with formal markers. Markers degrade over time, so that by the time the graveyard census was taken in 2003-2004, many graves could not documented because their information is illegible. Other markers were destroyed entirely in the period between the burial and the time the GCD was collected.

The RMS reflects the experiences reported in 1,396 households but omits the experiences of nearly 190,000 households not sampled. The HRVD reflects the experience of 7,668

respondents, but approximately 940,000 other East Timorese did not give testimonies to CAVR. However, even if the HRVD and RMS did reflect the experience of every living person in East Timor, many deaths would still remain undocumented because all the people who could remember them have died, left the country, or were psychologically or physically unable to recount the stories during the data collection period. In villages where mortality was especially heavy, there may have been no witnesses who survived until 2002-2003. Other families may have left Timor entirely, taking with them their social memory of the deaths. Still other families may have decided to keep secret their past experiences, so it may not be possible to directly document deaths in their family. Social memory is always partial.

The RMS uses standard household survey techniques based on the reported deaths to estimate total number of killings and deaths by hunger and illness. However, these totals are estimates of the total number of deaths that were possible to be remembered by current residents of Timor-Leste, which is a subset of the total deaths that actually happened. The ratio of the deaths remembered by current to all deaths is called the coverage rate.

An alternative method to estimate the total deaths uses multiple systems estimation (MSE). This method is used to correct censuses by comparing coverage among different documentation projects. MSE estimates of the number of deaths due to hunger and illness and due to killings are presented and compared to the RMS estimates. For killings, the MSE estimates are recommended, while for the estimated total number of deaths due to hunger and illness, the RMS estimates are recommended.[17]

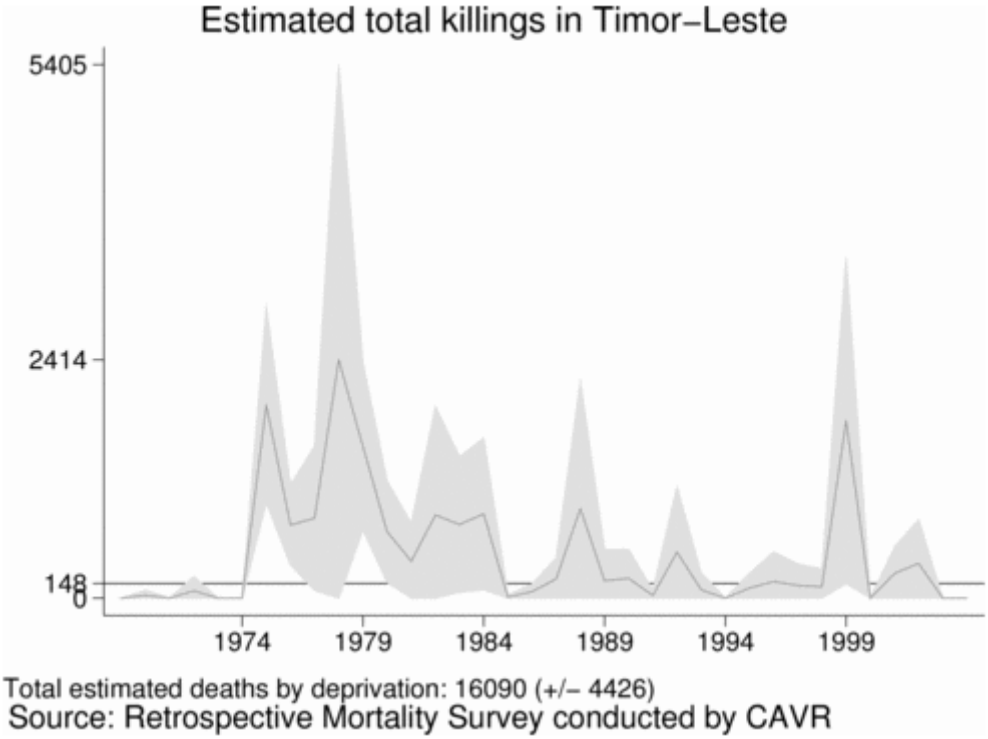
1.4 Estimates of Killings, Deaths due to Hunger and Illness, and Displacement

1.4.1 Killings

The annual total number of killings can be estimated from the RMS, and the results are presented below in Figure 1. This figure follows relatively high levels of killings in the 1975-1979 period, with additional peaks in the early 1980s and a spike in 1999. There are 16,000 total killings estimated by the RMS, with a margin of error of +/- 4,400. The RMS estimate of killings is based on only 235 reported killings. Consequently, the error is substantial, as seen by the many years for which the error bands touch zero. For these years, the hypothesis that the estimated number of deaths is zero cannot be rejected. Furthermore, the error bands are sufficiently wide that many different patterns could be possible.

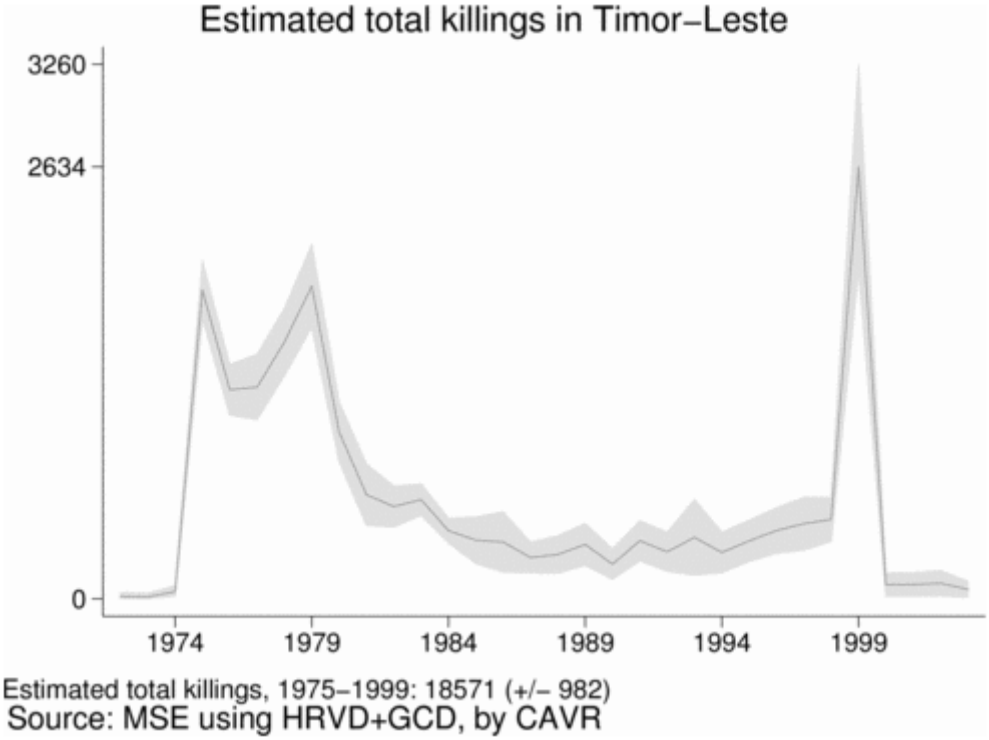
[17] MSE is widely used in estimating the under-reporting levels in population censuses . See for example Robinson J G, Ahmed B, das Gupta P and Woodrow K (1992), Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis, Journal of the American Statistical Association, 88(423), p1061-1071.

Figure 1



As with deaths due to hunger and illness, it is possible to estimate the annual total number of killings using multiple systems estimation. For killings, the CAVR testimonies (denoted HRVD) document nearly one-third of the total estimated killings, whereas as mentioned above, there are only 235 documented killings in the RMS. Consequently, the MSE for killings combines the GCD and the HRVD data. The results are shown in Figure 2. The MSE estimates 18,600 total killings (+/- 1,000). The vertical axis notes the maximum of the error (3,260) and the maximum estimated value (2,634), both of which occur in 1999. The pattern over time is much clearer in the MSE than in the survey estimate: the estimated total number of killings rises from nearly zero in the pre-invasion period to peaks in 1975 and 1979. Killings decline thereafter and through the 1980s and 1990s. A spike in 1999 marks the high point of estimated killings, significantly greater than any other year.

Figure 2



The killing estimates suffer from an important lacuna in the data: 1991 should have a small peak representing the Santa Cruz massacre, but insufficient reports of the event were captured to estimate the killings in that year correctly. The estimate for 1991 is actually slightly lower than the estimate for 1990. This illustrates a problem with all large-scale data collection: killings are relatively rare events across the entire population, so probability samples are unlikely to capture specific events (e.g., there were no reports of Santa Cruz in the RMS). In qualitative reports such as the HRVD, investigative resources must be specifically devoted to specific events, or there is no guarantee the events will be documented. Although the HRVD received more than twenty reported deaths for this event, this is only a small fraction of the total. Much smaller events were more thoroughly covered, and so the estimates over time do not accurately reflect the importance of this year.[18]

Unlike for deaths due to hunger and illness, killings are unlikely to be substantially underestimated. First, killings are less likely to affect entire families than hunger and illness deaths, so there are more likely to be surviving relatives to report these events. Second, the ratio of documented killings to estimated killings (the coverage rate) is 0.637, which is higher than the ratio of documented hunger and illness deaths to estimated hunger and illness deaths (0.513). The higher coverage rate for killings means that MSE itself could correct better for the unreported killings than for the unreported deaths due to hunger and illness. Furthermore, killings are relatively rare, and so the kind of speculative analysis using census-based crude

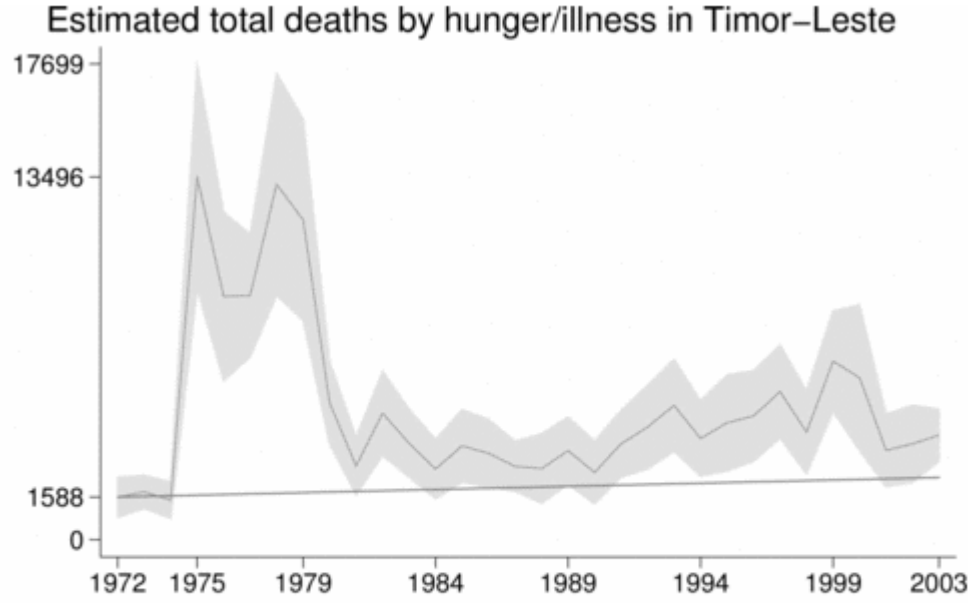
[18] Note that the margin of error specifically represents these "holes" in data of this kind. That is, by its nature sampling only captures information on a small number of the total events. Some large events (such as the Santa Cruz massacre) may be missed. The estimated error of the estimated total number of events is designed to reflect the uncertainty around the estimated total, including the fact that some large events may be missed.

death rates conducted for deaths due to hunger and illness is not possible for killings. Given this argument, we recommend that the finding for killings be that there were approximately 18,600 killings, with a margin of error +/- 1000.

1.4.2 Deaths due to hunger and illness

The analysis of the total number of deaths by hunger and illness begins with an analysis of the total estimated deaths from the RMS, and the results are presented below in Figure 3.[19] The estimated deaths are presented against a baseline of deaths projected from the 1972-1974 death rates due to hunger and illness (described below). The number of deaths attributed by respondents to "hunger or illness" rises to its highest levels during the immediate post-invasion period, 1975-1980. During the period 1983-1998, the estimated total fluctuates around a median of 3,632 estimated annual deaths. The annual total rises slowly during the 1990s, reaching a final peak in 1999.[20]

Figure 3



The total estimated number of deaths due to hunger and illness 1975-1999 is approximately 143,700, with a margin of error of approximately +/- 11,000.[21] Some of these deaths are natural in the sense that they would occur in the absence of conflict or famine. An assessment of deaths

[19] The data are insufficient for a three-system estimation of the deaths by hunger and illness due to the low coverage rate of deaths by hunger and illness in the HRVD.

[20] As is discussed in the methodological section, the slow rise in estimated deaths due to hunger and illness from the early 1980s through the late 1990s is a consequence of increasing population and a decrease in the number of deaths that are lost because no relatives survived until 2004 to be surveyed.

[21] The total margin of error is calculated by taking the square root of the sum of the variances of the annual estimates and multiplying by the conventional 1.96 to create a 95% confidence interval.

that could be attributed to the conflict must consider first how many deaths would likely have occurred as a result of hunger and illness in the absence of the conflict.

The immediate pre-invasion years 1972-1974 provide a peacetime baseline of natural deaths due to hunger and illness. First, to create a baseline population, population estimates for 1971-2003 were interpolated between the total reported in the 1970 Portuguese census (609,477) and the 2004 census total (924,642). From the RMS, the estimated number of deaths due to hunger and illness in 1972-1974 was approximately 1686-2252. Death rates for each year were computed by dividing the projected population for each year by the RMS estimated number of deaths for that year. The average of the 1972-1974 annual death rates due to hunger and illness was calculated.

The estimated death rate (3.1 deaths per 1,000 people) was applied to the projected population for each year through 2003.[22] These estimated death rates form the baseline shown in Figure 3. By subtracting the peacetime baseline projection from the annual RMS estimate, there are 84,200 deaths due to hunger and illness (+/- 11,000).[23] These deaths constitute our estimate from the RMS of the deaths which exceed the total that would be expected if the death rate due to hunger and illness had continued as it was in the pre-invasion peacetime period.

An alternative method to estimate the approximate total number of excess deaths is to calculate a two-system MSE using the RMS and GCD. This should be approached with caution, for two reasons. First, the number of reported hunger and illness deaths in the RMS in 1974-1999 (2,231) is small relative to the RMS total estimate of deaths due to hunger and illness (143,700). Second, two-system estimates can be biased as a result of uncontrolled correlation between the systems.[24] For example, if deaths that were unlikely to be buried in a public graveyards were also unlikely to be remembered by survey respondents, then the two systems would have a positive correlation. Note that this correlation is likely in the immediate post-invasion years when many people were living in very difficult conditions, at first moving constantly and later being held in internment camps. Both conditions would tend to lead both to catastrophic deaths of entire groups and to situations in which relatively few people were buried in public graveyards with permanent markers. The positive correlation between the GCD and RMS in extraordinary years would bias the two-system estimate downward, potentially significantly.

The MSE estimates are shown in Figure 4. The total estimated deaths by hunger and illness in 1975-1999 is approximately 123,500 (+/- 5,200).[25] The estimated deaths in excess of what would

[22] This baseline rate (3.1 deaths per 1000 people) is low: the regional average for Southeast Asia in the early 1970s was 12-14 per 1000. The methodological appendix addresses how the underestimates could be adjusted using census-based estimates of the crude death rate (CDR). However, adding the census-based measures requires many assumptions about the quality of the CDR estimates. The core findings presented here are based only on the 1970 and 2004 census estimates and the data collected by the CAVR.

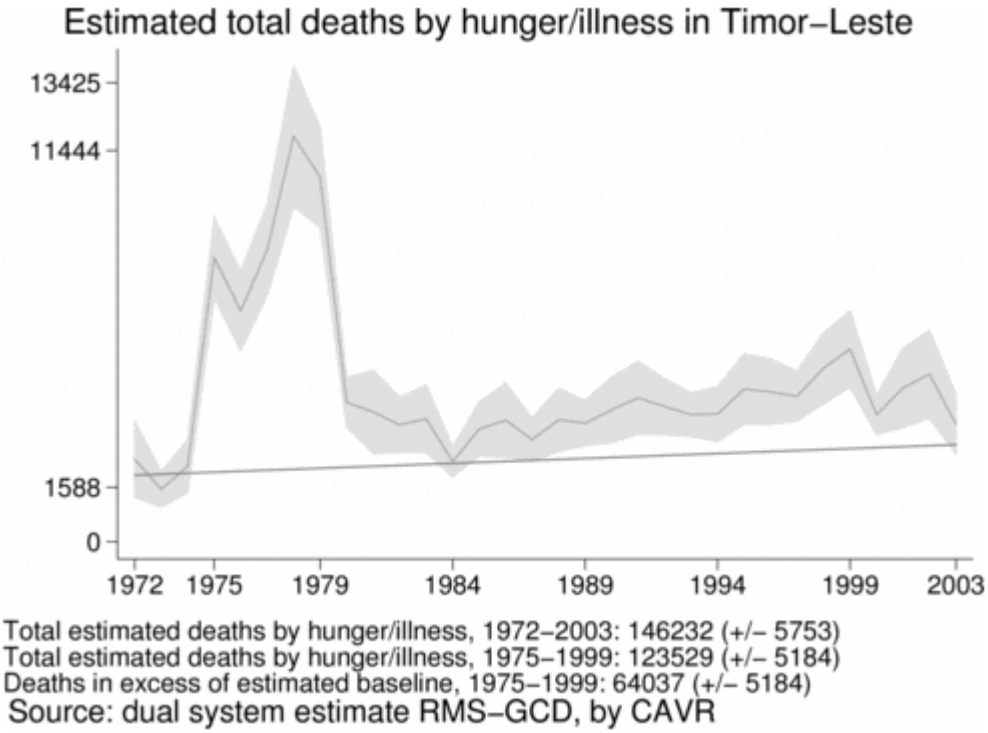
[23] The margin of error of the excess deaths is calculated in the same way as the total margin of error, including the standard error only of the years which contribute to the excess total.

[24] As mentioned earlier, the data are inadequate for three-system models of deaths due to hunger and illness. There are 5101 deaths due to hunger and illness reported in the HRVD, approximately twice as many as in the RMS, but this is nonetheless a small fraction of the expected total number of deaths due to hunger and illness. The RMS is used in preference to the HRVD because the RMS was collected by a probability sample. The HRVD is a convenience sample, and using it would require the assumption that all deaths in each year had the same probability of being documented. See the methodological appendix for an explanation of how the MSE was calculated.

[25] For the MSE, only the named deaths reported in the RMS are included. The RMS sampling weights were not used. The GCD does not include a manner of death, so records from the GCD are allocated as described in the methodological appendix.

be expected by the peacetime baseline is 75,000 (+/-5,200). The MSE estimate is lower than the survey estimate, which is consistent with the hypothesis that there is positive correlation between the RMS and the GCD. The median number of deaths due to hunger and illness during "normal" occupation years 1983-1998 estimated by MSE (3,727) is similar to the value found by the survey (3,632). That is, the survey estimates and the MSE estimates for "normal" occupation years are similar, but the RMS has higher estimates for the extraordinary years. It is likely that during normal years, relatively fewer deaths occur in catastrophic events that eliminate entire families (causing survey underreporting), and relatively fewer deaths are left outside public graveyards. Consequently, during normal years, the MSE and the survey provide similar estimates. During extraordinary years, both methods underestimate total deaths due to hunger and illness, but the MSE underestimates slightly more because a small proportion of all deaths are buried in public graveyards during extraordinary years.

Figure 4



The pattern shown in Figure 4 was similar to the RMS estimate, but the magnitude is lower: note that the maximum estimated value in Figure 4 is 11,444 whereas in the RMS estimate it is 13,496. There was one difference in the patterns over time: the MSE estimated totals for 1975 were lower than the estimates for 1978 and 1979, whereas in the survey, the estimates for the three years were close together. In the RMS estimate, the error bands for the three years were large relative to the differences among the years, and therefore, the hypothesis cannot be rejected that the real totals in those years are equal.

The MSE estimates suggest that 1975 had fewer deaths than implied in the RMS estimates. Other than this difference, the MSE and RMS estimates are similar in magnitude and pattern. It should be emphasized that the two estimates are methodologically very different: magnitude in the RMS was driven by the survey weights, while the preponderance of the data in the MSE came from the

number of graves with dates in each year and the matching of the deaths identified in the RMS to the names, dates and locations of graves.

Combining the results from the two estimates, a highly conservative estimated minimum number of deaths by hunger and illness in excess of the peacetime baseline is between 75,000 and 86,500. These estimates draw only on the 1970 and 2004 censuses and the CAVR's own data. These estimates should be explicitly understood as estimates of the total deaths due to hunger and illness which were possible to remember in 2004. This is a significant limitation on the calculations which can be made. Both the RMS and MSE estimates are substantially conservative because many deaths could not be remembered by 2004. Some deaths left no surviving family members available to report the death in 2004, and some deaths during extraordinary years were not buried in public graveyards. The years in which the survey is most likely affected by the loss of entire families are also the years in which people are least likely to be buried in public cemeteries. This positive correlation between the GCD and RMS data creates an underestimation in the MSE.

In the methodological appendix, a model is presented for adjusting the RMS and MSE estimates to correct for the loss over time of knowledge about deaths. This model uses additional census information, including crude death rates estimated by the US and Indonesian governments. If the assumptions in the underlying data and in the models were correct, the total deaths due to hunger and illness in excess of the peacetime baseline could be 103,000, with a possible (but improbable) high-end estimate of 183,300. Given the uncertainty in these models, we recommend that the finding be that at minimum, during the period 1975-1999 100,000 people died due to hunger and illness in excess of the peacetime baseline.

1.4.3 Displacement

The core findings from the analysis of displacement are outlined below:

- Displacement was widespread: 55.5% of surveyed households[26] reported one or more displacement events, for a total of 2011 reported displacement events between 1974 and 1999.
- When projected to the total population, the surveyed results represent 108,200[27] displaced households experiencing 282,800[28] displacement events.
- Most displacements were local. Of all displacement events, 54.3% are within subdistrict, 15.6% are within district, 17.4% are within region, 9.3%% are within East Timor, and 2.4% are outside of Timor.[29] However, in 1999, the displacements that take the household out of East Timor increase to 19.3% (+/-6.1%) of displacements in that period.
- Many displacements occurred in rapid succession: 22.2% of displacement events lasted one month or less, and 50.1% lasted one year or less. However, other displacements were

[26] According to the 1990 census, there are approximately 4.5 people per household. This figure increases to 4.75 people per household (924,642/194,943) in the 2004 census. The nominal confidence interval is 51.8%-59.2% of households.
[27] The nominal confidence interval is 101,013-115,412 households.
[28] The nominal confidence interval is 251,631-313,990 events.
[29] The nominal margin of error is +/- 10.4% for within subdistrict, and 4.6% or less for the other estimates. This finding may be limited by the restriction that people in refugee camps in West Timor were not interviewed.

very long, so that the mean displacement period lasted 46.7 months.[30] Tens of thousands of households considered themselves displaced from the immediate post-invasion period in 1975q4 continuously until 1999q4 (see Figure 6 for more analysis).

- The institution that respondents reported most frequently as the group telling them to move was the Indonesian military (46.4%), followed by FALINTIL (15.0%) and militias (8.8%).[31] Respondents reported that "conflict" motivated 52.3% of their displacements, with "forced by Indonesian military" contributing an additional 16.3%.
- Most displacements occurred in 1975-1980. The maximum years are 1975 and 1976, with 61,400 (+/- 13,300) and 59,800 (+/- 7,200) displacement events, respectively. The events of 1999 were substantially fewer, with approximately 28,100 (+/- 5,600) events.

Figure 5

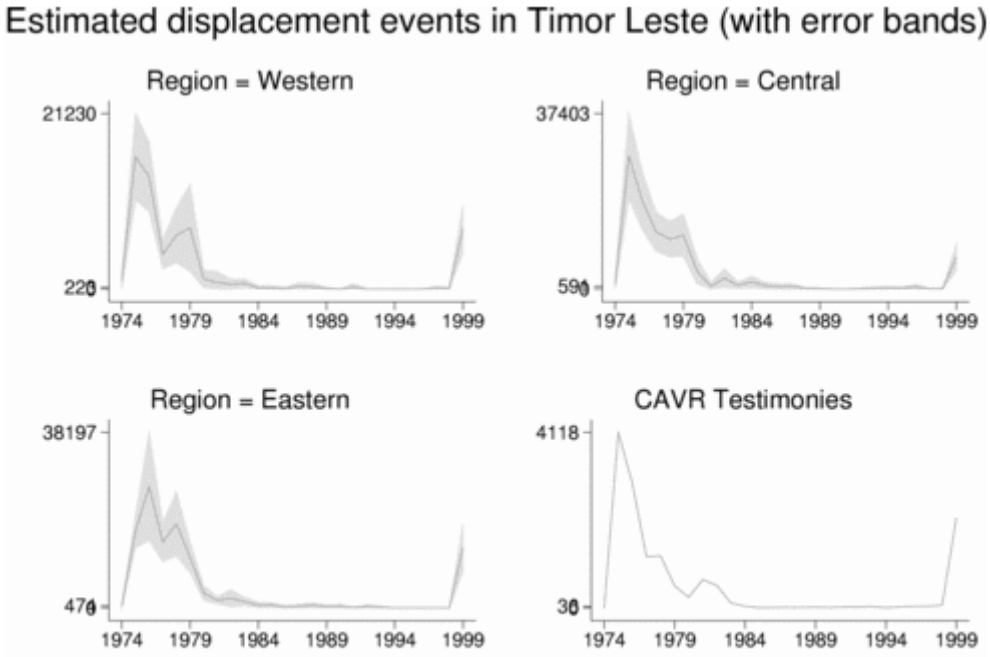
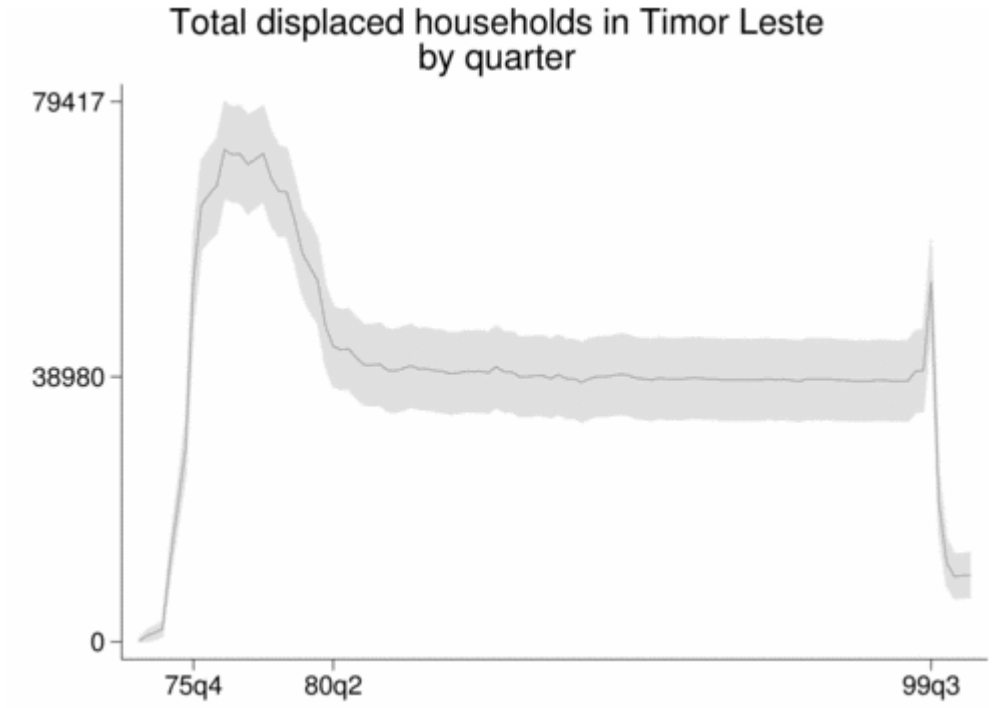


Figure 5 shows the number of displacement events by year in three regions of East Timor, as well as the total displacement events reported to the CAVR in qualitative testimonies. The vertical axes label the maximum values (at the top of the error band) for each region. Figure 5 shows 1975 and 1976 as the peak years in overall displacements, with 1977-1979 and 1999 at roughly equal lower levels. We note that 1975 and 1976 cannot be statistically distinguished from each other in any of the regional estimates; in formal language, we cannot reject the hypothesis that these years had equal numbers of displacement events. The Western and Central regions show a decline from the 1975-76 peaks to lower levels through 1979; the decline is then toward zero. In contrast, the Eastern region shows displacement levels in 1979 that are nearly the equal of the displacement intensity of the immediate-post-invasion period.

[30] The nominal confidence interval is 41 – 52 months.
[31] The nominal confidence interval is 41 – 52 months.

Another way to look at displacements is to consider how many households are displaced during each period. Figure 5 showed how many households were forced to move during each year. Figure 6 shows how many households were forced to live in a place they did not consider their home during each period. That is, Figure 5 shows how many "displacement events" were suffered in each period, while Figure 6 shows how many households are in the status of "displacement" in each period. The vertical axis notes the maximum value at the top of the error band and the median value from the third quarter of 1980 (1980q3) through 1999q2.

Figure 6



Source: Retrospective Mortality Survey (Displacements Module), CAVR

People were displaced in late 1975 and early 1976, and they were unable to return to their homes for a long period. Additional displacements continued in 1977, with a few more in 1978, and people displaced earlier are still away from their homes. Not until 1979 do large numbers of households settle in places they consider "home." From early 1980 until 1999, an estimated 39,000 households continue to consider themselves displaced. An additional 15,000 households are displaced in the third quarter of 1999. However, in the fourth quarter, we estimate that more than 32,500 households returned to their homes, and the estimated number of displaced households dropped to approximately 20,400; to 11,700 in 2000q1, and to 9,600 in 2000q2.

1.5 Descriptive statistical Analysis of Fatal Violations Reported to the Commission

This section describes the pattern of killings and disappearances reported to the Commission in narrative statement taking process (the HRVD). The magnitude and patterns described here do not represent the total magnitude and overall pattern of killings and disappearances. Rather this

analysis describes the pattern and trend of killings and disappearances which is known through the Commission' s qualitative statements.[32]

1.5.1 The Reported Pattern of Killings and Disappearances of Non-Combatants Over Time

The pattern of reported killings and disappearances varied substantially over time. As can be seen in Figure 7, 67.4% (3,451/5,120) of reported killings are concentrated in the period 1975-1981. 16.4% (838/5,120) of reported killings occurred during the UN-administered Popular Consultation in 1999. The highest counts of killings reported to the Commission were during the period of the invasion by the Indonesian military and initial years of occupation. Although the year with the highest reported counts of non-combatant killings was 1975, the open-ended nature of the narrative statement taking process was such that a considerable amount of date imprecision was encountered in statements which reported killings in the late 1970's.[33] It is therefore likely that some of the non-combatant killings which were reported to have occurred in 1975 may have actually occurred in 1976 or 1977.

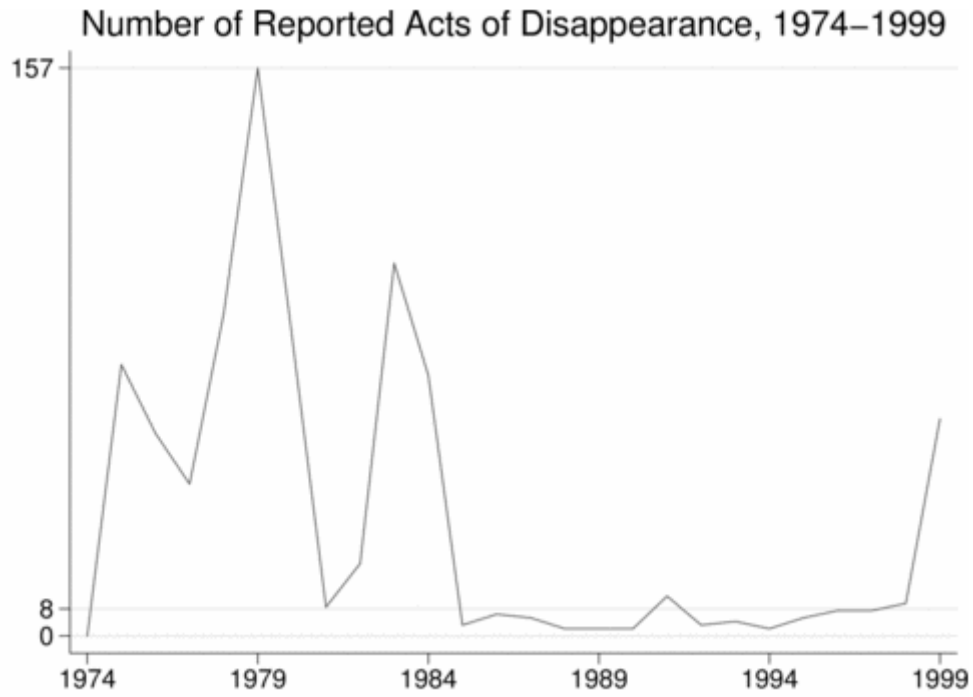
Figure 7

The counts of disappearances reported to the Commission are substantially lower than that of reported non-combatant killings: 5,120 non-combatant killings were reported to the Commission, whereas 835 disappearances were reported to the Commission. Furthermore, the reported pattern

[32] See Section 2.3.1.3 for a detailed discussion about the nature and limitations of data collected through the Commission's statement-taking process.
[33] See Section 5.4.2.3.1 of the Statistical Methodological Appendix for more detailed description of date imprecision in the statement-taking process.

of disappearances is substantially different than that of non-combatant killings, as can be seen in Figure 8. Whereas large-scale non-combatant killings were overwhelmingly concentrated in the initial invasion years, large-scale disappearances were mostly concentrated towards the end of the initial invasion period in 1979 and at the start of the normalization and consolidation period of the Indonesian occupation around 1983/1984: 40.0% (332/835) of individual disappearances reported to the Commission occurred either in 1979, 1983 or 1984. The reported pattern of disappearances and non-combatant killings is consistent with the hypothesis that the two violations phenomena were driven by different policies or practices of those responsible. In particular, disappearances appear to have been used in a more targeted fashion as a counter-resistance tool by the Indonesian military.

Figure 8



Source: Database of Narrative Statements Given to the CAVR

20.9% (1,070/5,120) of killings documented by the Commission's statement-taking process occurred in 1975. As shown in Table 1, of the documented killings in 1975, 26.5% (283/1,070) of these killings do not contain information about the month in which they occurred. Whereas, 19.5% (348/1,070) of these killings occurred during the time of the internal party conflict, and 32.7% (350/1,070) occurred in December at the time of the launch of the Indonesian military invasion of Timor-Leste.

Table 1: Count of Reported Acts of Civilian Killing, 1975

Year	Frequency	%
Jan	4	0.4
Feb	6	0.6
Mar	6	0.6
Apr	20	1.9
May	3	0.3
Jun	3	0.3
Jul	6	0.6
Aug	194	18.1
Sep	154	14.4
Oct	30	2.8
Nov	11	1
Dec	350	32.7
Not Reported	283	26.5
Total	**1,070**	**100.0**

Source: Database of Narrative Statements Given to the CAVR

It is notable that only 3.8% (41/1,070) of documented killings in 1975 occur in October and November. Hence, the Commission's statement-taking process is consistent with the hypothesis that large-scale killings occurred during the internal party conflict in August and September, then there was a relative lull in violence in the form of killings prior to large-scale in December at the time of the Indonesian military's invasion of Timor-Leste.

1.5.2 The Reported Pattern of Killings and Disappearances of Non-Combatants Over Space

Data based on convenience samples cannot be used to assess directly the differences in the magnitude of violations between regions and districts. Furthermore, such data are representative only of the total extent of violence from region to region insofar as the deponents whose statements were taken are representative of their local population and were selected in proportion to the violence suffered in each district.

Figure 9 shows the counts of reported killings and disappearances by district in which the violation occurred, as reported in the Commission's statement-taking process. Ermera has substantially more reported killings than any other district, accounting for 18% (920/5,120) of all reported killings. Relatively few non-combatant killings in Indonesia, Dili and Liquiça were reported to Commission.

Figure 9

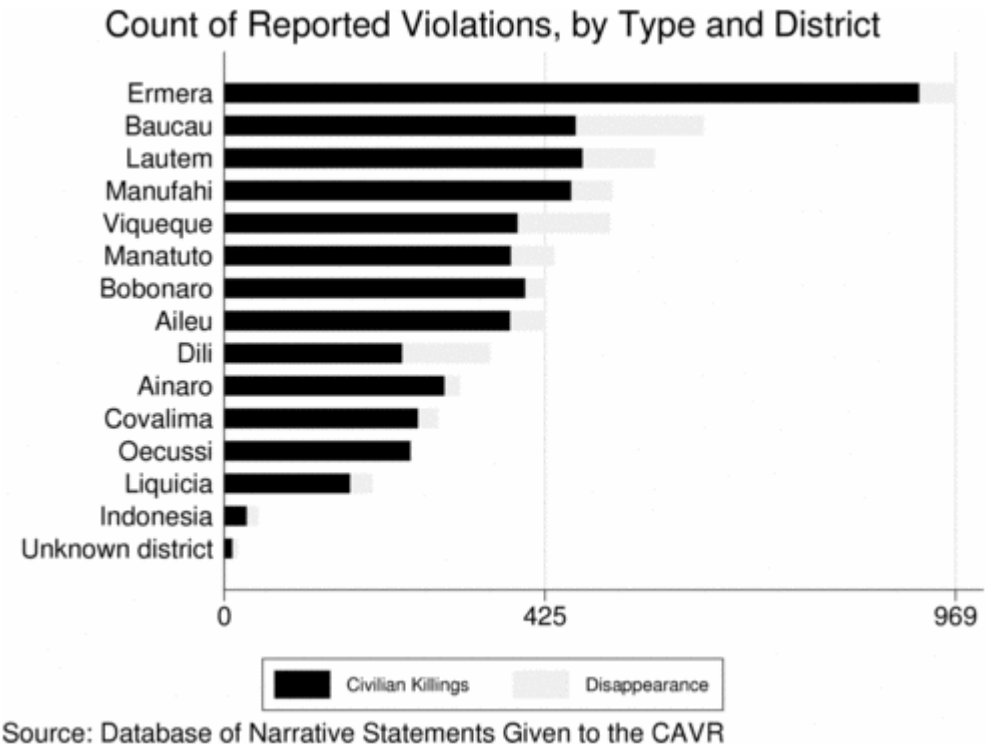
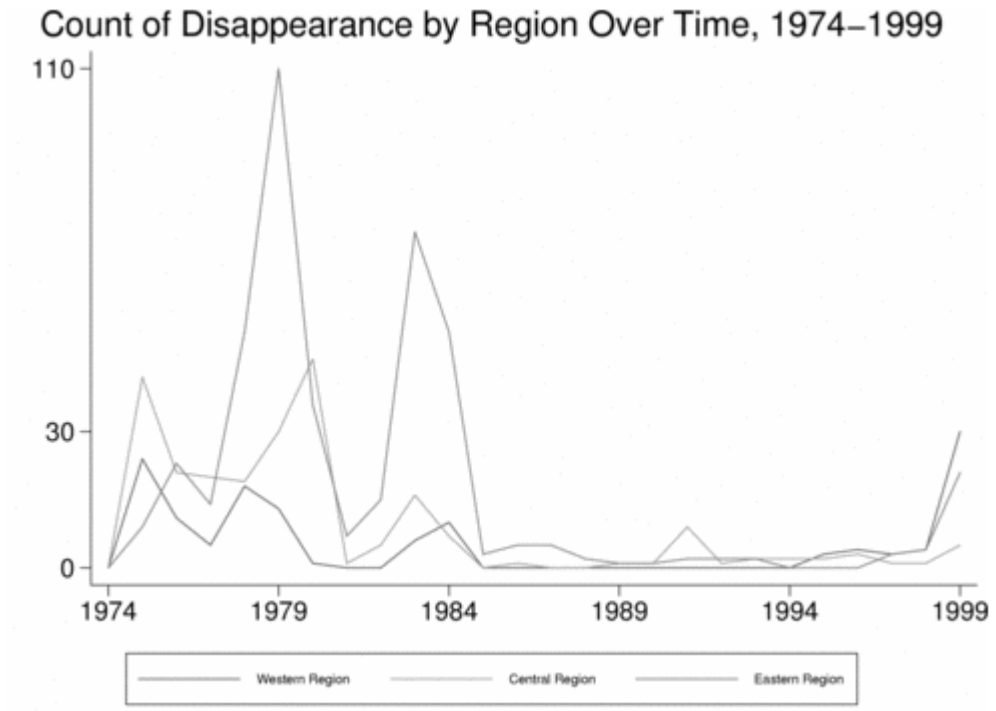


Figure 10 shows that reported disappearances were concentrated mainly in the Eastern and Central districts: in particular of the disappearances reported to the Commission, 20.2% (169/835) occurred in Baucau, 14.7% (123/835) in Viqueque, 13.9% (116/835) were in Dili and 11.4% (95/835) were in Lautem.

Figure 10

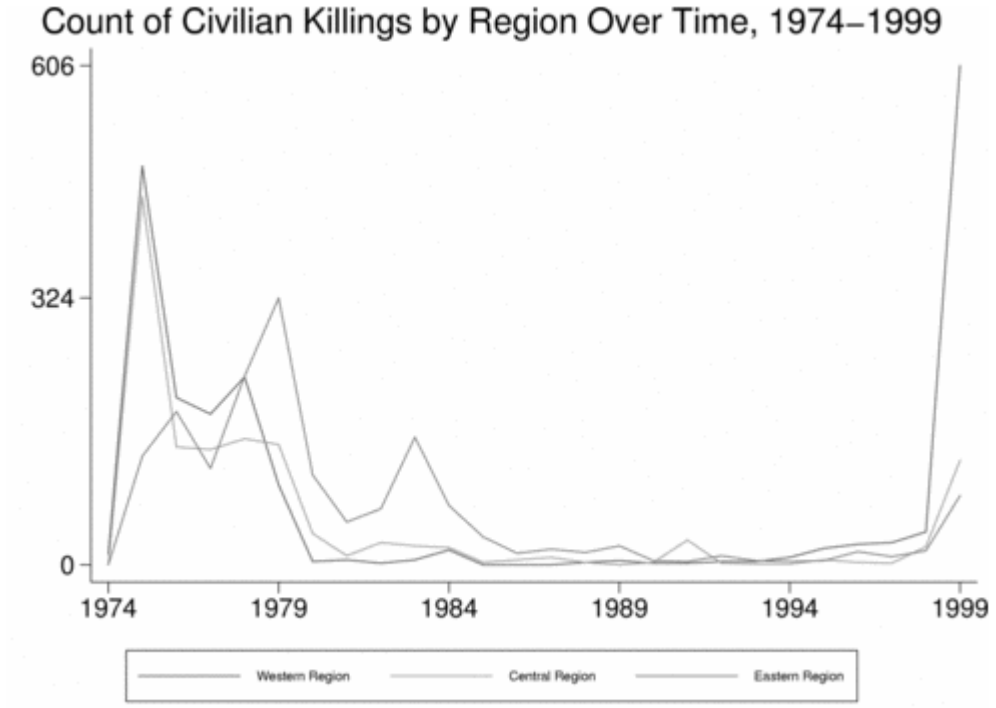


Source: Database of Narrative Statements Given to the CAVR

1.5.3 The Reported Pattern of Killings and Disappearances of Non-Combatants Over Time & Space

As Figure 11 shows, reported killings start in the western and central regions at the time of the initial Indonesian invasion. Then between 1978 and 1981, most reported non-combatant killings are then in the Eastern region and Central regions, with few reported non-combatant killings occurring in the Western region. In 1999, 72.3% of reported non-combatant killings occurred in the Western region. The Commission's narrative statement data is consistent with the hypothesis that, between 1975 and 1984 and again in 1999, large-scale individual non-combatant killings broadly tracked the movements across time and space of the invading Indonesian military.

Figure 11



Source: Database of Narrative Statements Given to the CAVR

The pattern of reported disappearances over time and space is notably different to that of non-combatant killings, as can be seen in Figure 10 above. Although some disappearances are reported around the time of the Indonesian invasion and again in 1999, disappearances do not appear to be associated with large-scale military operations in the same way non-combatant killings are. Rather, two periods of large-scale disappearances were reported: the first period occurring from 1978 to 1980 and the second period from 1983 to 1984. During the first period of large-scale disappearances, 60.2% (198/329) of reported disappearances were concentrated in the Eastern region, with 25.9% (95/329) occurring in the Central region and 10.0% (33/329) in the Western region and Indonesia. During the second period of large-scale disappearances, which occurred between 1983 and 1984, 72.0% (126/175) occurred in the Eastern region, 13.1% (23/175) in the Central region and 13.1% (23/175) in Indonesia and the Western region. This data on disappearances is consistent with the hypothesis that disappearances were used mostly in the Eastern districts as a counter-resistance tool against those suspected of being members or associates of the resistance groups.

1.5.4 The Pattern of Reported Killings and Disappearances of Non-Combatants against Individual and Group Victims

Some killings and disappearances were reported to the commission as being perpetrated against a lone individual, whereas some others were reported as being perpetrated against multiple individuals at the same time. Figure 12, and 13 show the distribution of violations by victim group size for killings and disappearances.[34]

[34] As is the case with reported violations against individual victims, violations against victims in groups can be reported by more than one deponent. We matched group victim records to identify duplicate reports of the same

Figure 12

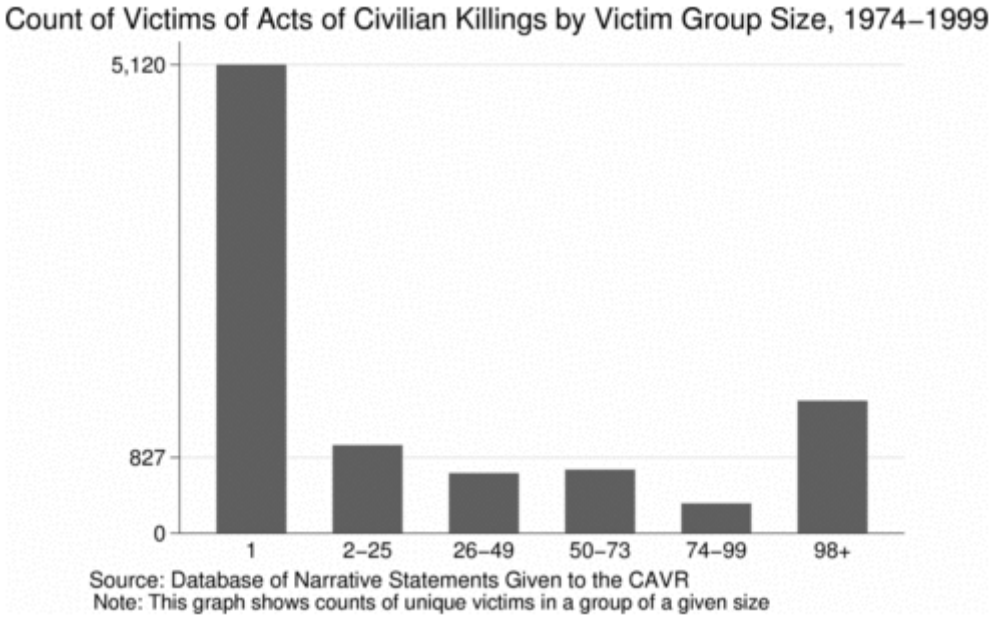
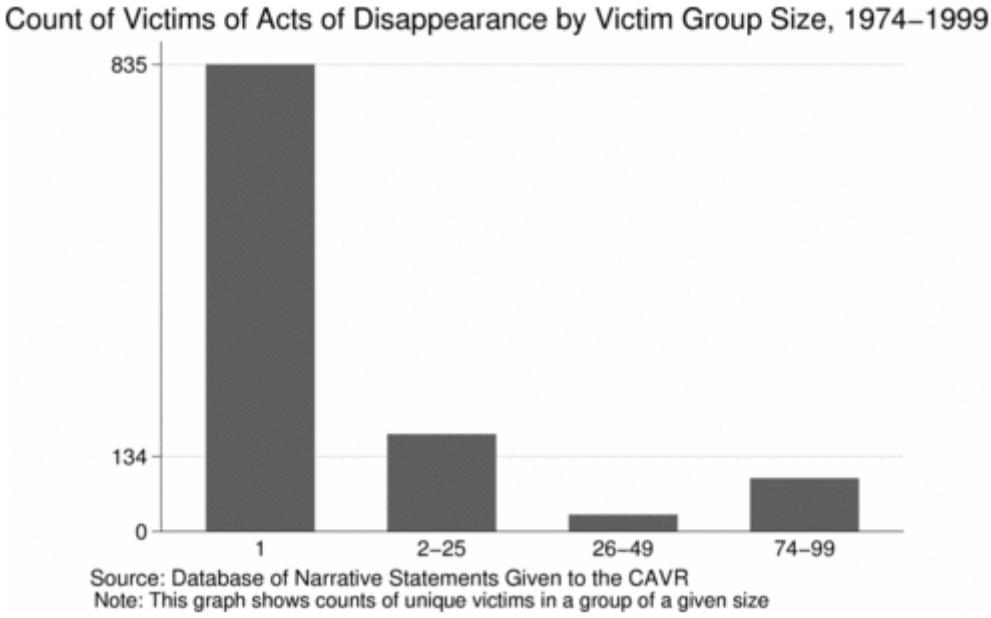


Figure 13



The nature of whether violations were committed against individuals one-by-one or collectively is similar for reported killings and disappearances. As Figure 12 and 13 show, 95.9% (5,120/5,339) of reported killings and 96.9% (835/862) of reported disappearances were perpetrated against people one-by-one. This empirical finding appears to be consistent with the hypothesis that the use of killings and disappearances as a form of oppression was used in a targeted fashion.

---

violation and victim in multiple statements. The methods used for matching are described in the Statistical Appendix in Section 3.?

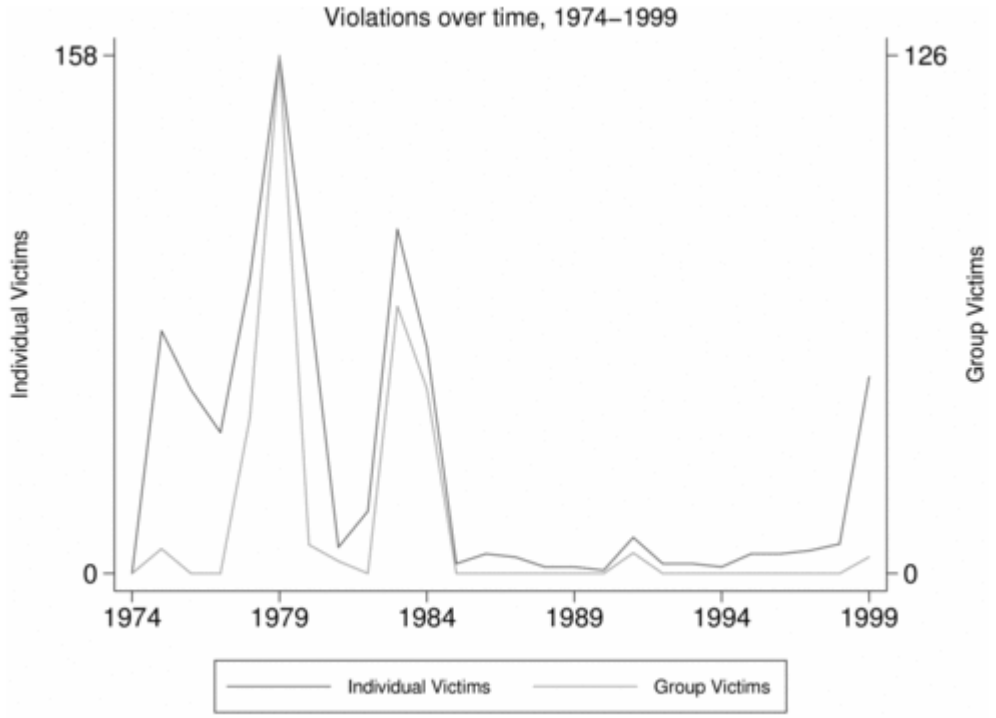
As Figures 14 and 15 show, in statements given to the Commission, both killings and disappearances against individual victims and group victims are positively correlated over time. When reported killings against individuals increase, so do reported killings against group victims. As is the case for reported disappearances.[35] Hence, large-scale reported group killings are concentrated in the early invasion years between 1975-1979, as are large-scale reported individual killings. Whereas large-scale reported group disappearances are concentrated during the period of counter-resistance campaigns in 1979 and 1984, as are large-scale reported individual disappearances.

Figure 14

Source: Database of Narrative Statements Given to the CAVR

[35] The correlation coefficient between reported individual killings and reported group killings over time is 0.95. Whilst the correlation coefficient between reported individual disappearances and reported group disappearances over time is 0.84.

Figure 15



Source: Database of Narrative Statements Given to the CAVR

There are at least two possible explanations for the respective positive correlation between (i) reported individual killings and group killings and (ii) reported individual disappearances and group disappearances:

i. individual killings and group killings are driven by either the same practices or policies of those responsible for these crimes, as is the case for disappearances., or

ii. others specifically identifying individual victims of killings and disappearances during large-scale military offensives in the late 1970's and between 1983-1984. Consequently, some deponents may have described these killings and disappearances as being suffered by anonymous groups of victims.

Nevertheless, whichever of these explanations is correct, the Commission's narrative statement data is consistent with the hypothesis that large-scale disappearances and large-scale killings were concentrated in time: large-scale killings being particularly concentrated in the early invasion and occupation years (1975-1979), whilst large-scale disappearances being concentrated in towards the end of the invasion years (1979) and during the counter-resistance crack-downs in the Eastern districts (1983-1984).

1.5.5 The Reported Pattern of Killings and Disappearances of Non-Combatants Across Demographic Characteristics and Political Affiliation of Victims

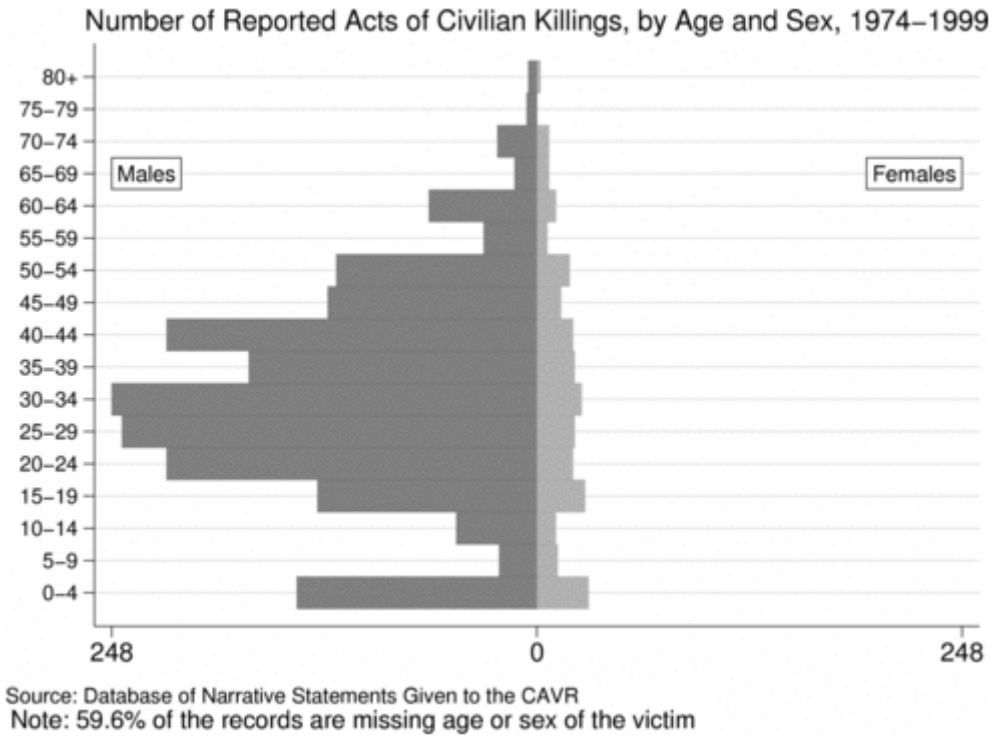
The Commission's research considered whether killings and disappearances were conducted in a systematic fashion and targeted at victims because of their demographic characteristics (such as age and sex) or particular political affiliations. This section describes the reported extent and

pattern of killings and disappearances against different sexes and ages of victims and against civilians, armed-resistance fighters and political activists.

Killings and disappearances reported during the Commission's statement-taking process were overwhelmingly against male victims. In particular, 86.9% (4,451/5,120) of reported killings were of male victims and 90.5% (756/836) of reported disappearances were against male victims. In addition, young adults between the ages of 20 and 34 were the most frequently reported age groups for victims of killings and disappearances: 37.4% (663/2,090) of killing victims whose ages were known were in this age group.[36] Whereas 40.0% (138/345) of disappearance victims whose ages were reported were between the ages of 20 and 34.[37]

As can be seen in Figure 16 and 17, young males between the ages of 20 and 34 were the most frequently reported victims of killings and disappearances to the Commission's statement-taking process.

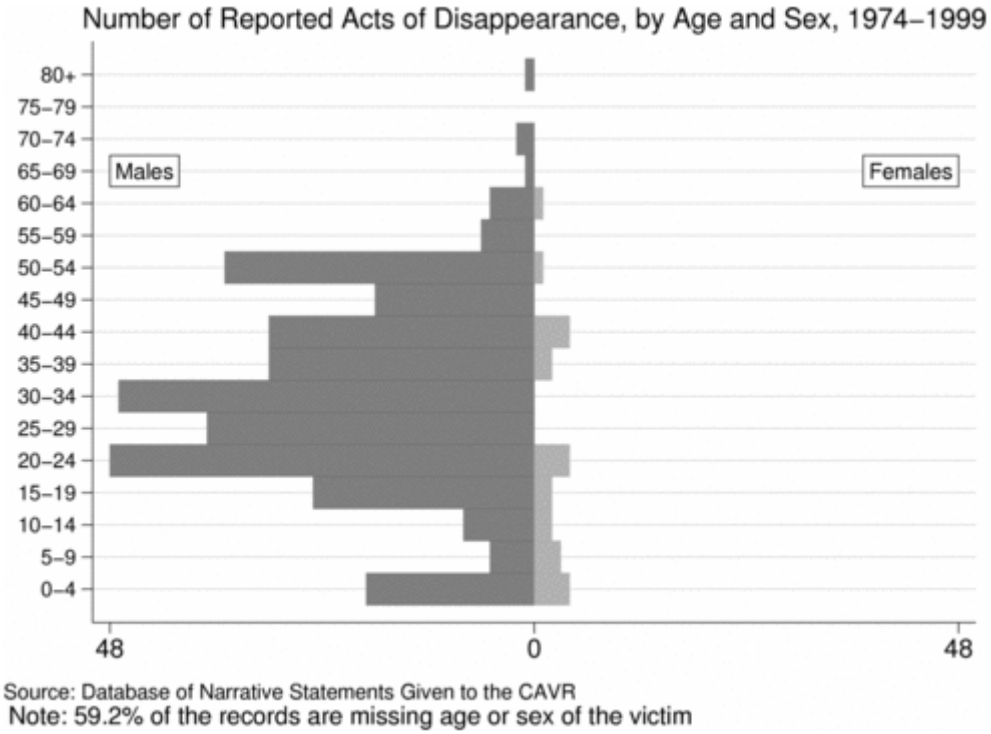
Figure 16



[36] Specific ages for 59.6% (3030/5120) of killing victims reported to the Commission's statement-taking process were either not known or not reported.

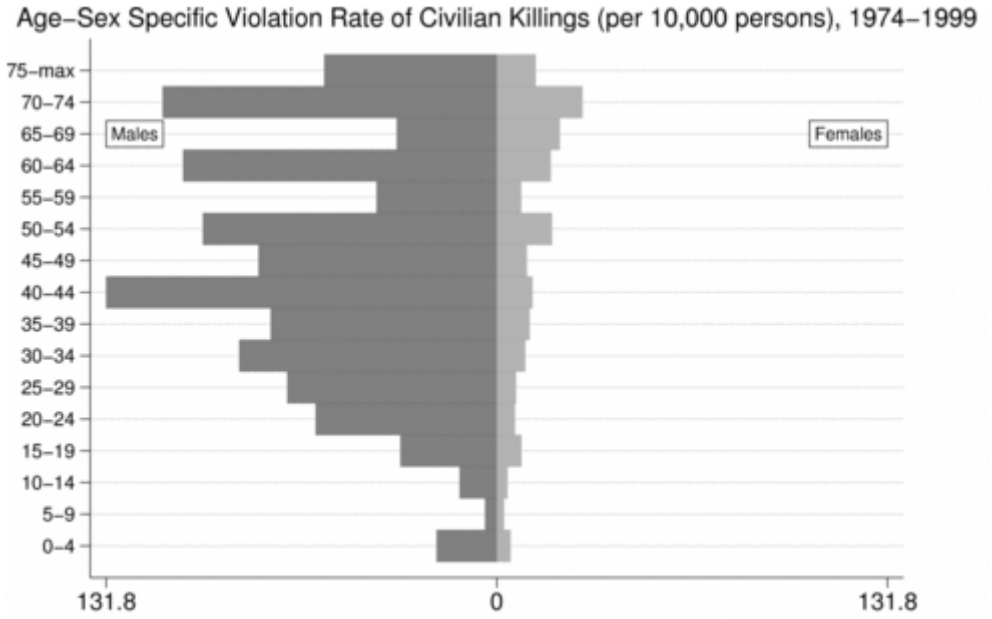
[37] Specific ages for 59.2% (490/835) of disappearance victims reported to the Commission's statement-taking process were either not known or not reported.

Figure 17



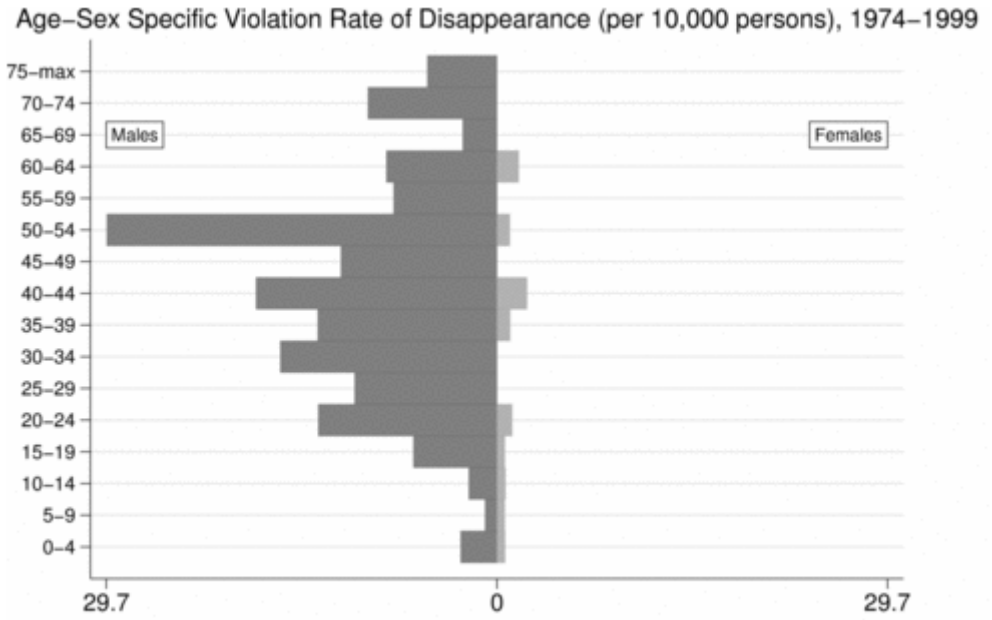
Moving from simple violation counts of killings and disappearances to population-based violation rates, notable differences can be observed. Relative to the overall Timorese population, middle-aged and elderly males experienced the highest rates (relative to their share of the population) of reported killings and males in the age-group 50-54 years-old experienced the highest rates (relative to their share of the population) of reported disappearances. These patterns are shown in Figures 18 and 19.

Figure 18



Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 59.6% of the violation records are missing age or sex of the victim

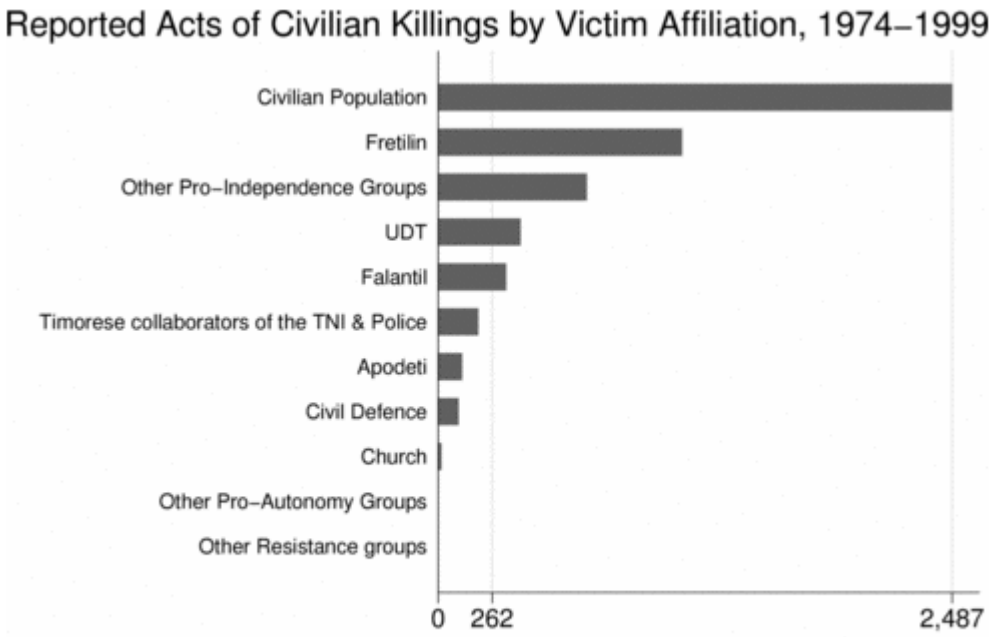
Figure 19



Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 59.2% of the violation records are missing age or sex of the victim

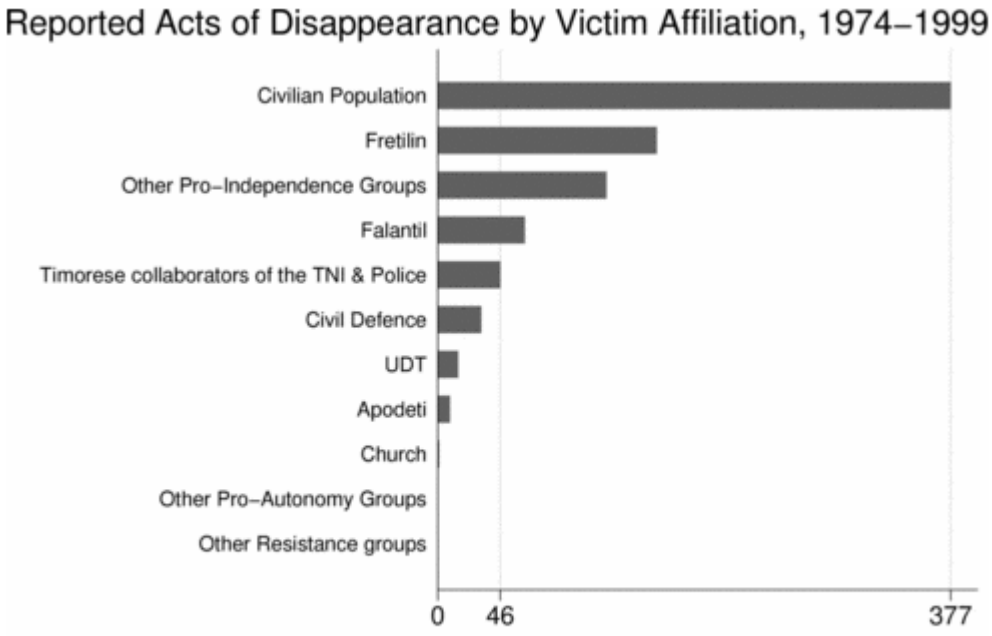
48.7% (2,487/5,120) of killings and 45.3% (377/835) of disappearances reported to the Commission were committed against the civilian population, including both those civilians who were not known to have a political affiliation and those who were formally part of a pro-independence group or political party, as can be seen in Figures 20 and 21.

Figure 20



Source: Database of Narrative Statements Given to the CAVR
Note: some violations may be counted more than once
because responsibility may be shared among perpetrators

Figure 21



Source: Database of Narrative Statements Given to the CAVR
Note: some violations may be counted more than once
because responsibility may be shared among perpetrators

40.9% (2,092/5,120) of killings reported to the Commission's statement-taking process were against victims who were either formally affiliated with Fretilin or a non-violent pro-

independence group. 6.4% (329/5,120) of reported killings were against victims who were reported to be affiliated with FALINTIL.

33.5% (280/835) of disappearances reported to the Commission's statement-taking process were against victims who were either formally affiliated with Fretilin or a non-violent pro-independence group. 7.6% (64/835) of reported disappearances were against victims who were reported to be affiliated with FALINTIL.

The Commission's narrative statement data is consistent with the hypothesis that the overwhelming majority of killings and disappearances were committed against members and suspected associates of the resistance movement and those persons who were not formally associated with a political party or armed group.

1.5.6 The Reported Pattern of Killings and Disappearances of Non-Combatants by Attributed Institutional Responsibility

A number of different institutions were involved in acts of killing and disappearance over the course of the conflict. The main institutional groups were the Indonesian military, FALINTIL, Timorese political parties (such as Fretilin, UDT and Apodeti), local administrative associates (such as Hansips and the Civil Defense Forces) and militias. This section reviews the reported levels of responsibility for killings and disappearances reported during the Commission's statement-taking process.

The majority of killings and disappearances reported to the Commission were attributed to the Indonesian military and their Timorese auxiliaries, as shown in Table 2: 57.6% (2,947/5,120) of the perpetrator involvement in fatal violations was attributed to the Indonesian military and Police, and 32.3% (1,654/5,120) to Timorese auxiliaries (such as the militias, civil defense force and local officials who worked under the Indonesian administration). In 29.6% (1,514/5,120) of reported killings and disappearances, institutional perpetrator responsibility was attributed to the resistance groups and pro-independence forces.

Table 2: Count of Reported Civilian Killings and Disappearances by Attributed Institutional Perpetrator, 1974-1999

Violation Type	Indonesian Military	Timorese Collaborators of TNI	Resistance Groups	Other	Civilian Population	Pro-Autonomy Groups	Unknown	Total
Civilian Killings	2,947	1,654	1,514	1,341	214	81	708	5,109
Disappearance	642	245	80	72	21	2	111	833
Total	**3,589**	**1,899**	**1,594**	**1,413**	**235**	**83**	**819**	5,942

Responsibility for violations may be shared among perpetrators, and therefore, columns may not be directly summed

Source: Database of Narrative Statements Given to the CAVR

The levels of attributed institutional responsibility for documented killings and disappearances varied over the course of the conflict. During the initial Indonesian invasion between 1975 and 1984, 62.3% (2,831/4,543) of documented killings and disappearances were attributed to the Indonesian military and police. Then during the period of normalization and consolidation of the Indonesian occupation, between 1985 and 1998, 64.6% (317/488) of documented killings and

disappearances were attributed to the Indonesian military and police. Then, in 1999 in the lead-up to and then aftermath of the UN-sponsored Popular Consultation, a significant structural change in the proportional responsibility of documented violations is observed. In particular, 9.5% (85/898) of killings and disappearances are attributed to the Indonesian military and police acting alone, 39.5% (355/898) are attributed to the Indonesian military and police acting in concert with the Timorese militias and 42.9% (385/898) are attributed to the Timorese militias acting alone.

In contrast, while 49.0% (561/1,145) of documented killings and disappearances in 1975 were attributed to Fretilin, 16.6% (563/3,398) of documented killings and disappearances between 1976 and 1984 were attributed to Fretilin.[38] Furthermore, 3.7% (18/488) of killings and disappearances between 1985 and 1998 were attributed to Fretilin and then in 1999 0.6% (5/898) of killings and disappearances were attributed to Fretilin. The Commission's narrative statement data is consistent with the hypothesis that most killings and disappearances attributed to Fretilin were committed during the 1975 internal party conflict.

1.5.7 The Association between Conflict-related deaths and Periods of Detention

The pattern of arbitrary detentions and civilian killings, which were reported to the Commission, are positively correlated over time.[39] In particular, both reported non-combatant killings and arbitrary detentions were overwhelmingly concentrated during the initial years of the Indonesian invasion and occupation as shown in Figures 22 and 23.

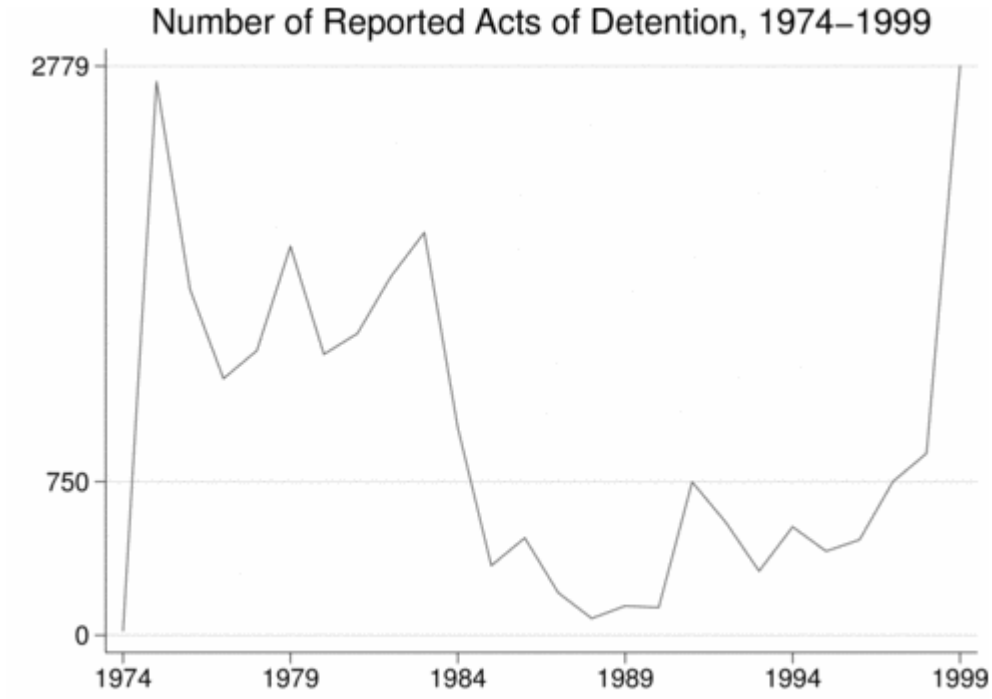
[38] Note that 42.2% (237/561) of documented killings and disappearances attributed to Fretilin in 1975 occurred during the internal political party conflict in August and September of that year. 4.8% (27/561) of documented killings and disappearances attributed to Fretilin in 1975 occurred in December, whereas 43.3% (243/561) of killings and disappearances in 1975 attributed to Fretilin did not contain specific information about the month in which the violation occurred.

[39] The correlation coefficient for the two series is 0.83

Figure 22

Source: Database of Narrative Statements Given to the CAVR

Figure 23



Source: Database of Narrative Statements Given to the CAVR

Furthermore, of those civilians who were reported to have died due to conflict-related reasons (namely either due to a civilian killing, death by hunger and illness, or disappearance), 98.6%

(10659/10809) were reported to have been arbitrarily detained at least once during the Commission's mandate period. As Table 3 shows, 3.5% (378/10,809) of these victims died whilst they were being held in detention.[40]

Table 3: Reported Fatal Violations & Their Detention Context by Geographic Location, 1974 - 1999

District	Victim detained at time of violation	Victim detained but unknown dates	Victim detained - but not at time % of violation	Victim never detained	Total
Lautem	49	186	365	16	**616**
% Lautem	8.0	30.2	59.3	2.6	100
Viqueque	26	222	945	24	**1,217**
% Viqueque	2.1	18.2	77.6	2.0	100
Baucau	40	151	887	15	**1,093**
% Baucau	3.7	13.8	81.2	1.4	100
Manatuto	13	104	1,069	1	**1,187**
% Manatuto	1.1	8.8	90.1	0.1	100
Manufahi	22	119	746	6	**893**
% Manufahi	2.5	13.3	83.5	0.7	100
Aileu	24	120	769	15	**928**
% Aileu	2.6	12.9	82.9	1.6	100
Ermera	51	107	1,506	32	**1,696**
% Ermera	3.0	6.3	88.8	1.9	100
Liquiça	21	37	379	6	**443**
% Liquiça	4.7	8.4	85.6	1.4	100
Dili	33	95	402	4	**534**
% Dili	6.2	17.8	75.3	0.7	100
Ainaro	16	58	270	1	**345**
% Ainaro	4.6	16.8	78.3	0.3	100
Covalima	12	23	445	1	**481**
% Covalima	2.5	4.8	92.5	0.2	100
Oecussi	27	13	201	16	**257**
% Oecussi	10.5	5.1	78.2	6.2	100
Bobonaro	40	72	926	12	**1,050**
% Bobonaro	3.8	6.9	88.2	1.1	100
Indonesia	4	7	44	1	**56**
% Indonesia	7.1	12.5	78.6	1.8	100
Unknown District	0	0	13	0	**13**
% Unknown District	0.0	0.0	100.0	0.0	100
Total	**378**	**1314**	**8,967**	**150**	**10,809**

Source: Database of Amnesty International Reports on East Timor

[40] For 12.2% (1,314/10,809) of conflict-related deaths suffered by individuals who were also arbitrarily detained during the Commission's mandate period, the dates of their detention was not known. Hence the Commission was not able to discern whether or not these deaths occurred whilst the individual w as being detained.

Of the fatal violations reported to the Commission, which occurred whilst the victim was being held in detention, 96.6% (365/378) were civilian killings or disappearances and 3.4% (13/378) were deaths due to hunger and illness. Whereas, the distribution by cause of death for those individuals who died of conflict-related causes outside of detention was substantially different: 49.0% (4,390/8,967) of the victims documented by the Commission died as a result of killings or disappearances whereas the remaining 51.0% (4,577/8,967) were deaths due to hunger and illness. Hence, the Commission's quantitative data is consistent with the hypothesis that detainees were at a relatively higher risk of being killed or disappeared whilst being held in-detention than when they were not being detained.

Table 4: Reported Fatal Violations & Their Detention Context by Fatal Violation Type, 1974 - 1999

	Civilian Killings	% Civilian Killings	Death due to Deprivation	% Death due to Deprivation	Disap-pearance	% Disap-pearance	Total
Victim detained at time of violation	295	5.8	13	0.3	70	8.4	**378**
Victim detained but unknown dates	807	15.8	273	5.6	234	28.1	**1,314**
Victim detained - but not at time of violation	3,887	76.1	4,577	94.0	503	60.4	**8,967**
Victim never detained	118	2.3	6	0.1	26	3.1	**150**
Total	**5,107**	**100.0**	**4869**	**100.0**	**833**	**100.0**	**10,809**

Source: Database of Amnesty International Reports on East Timor

The pattern of conflict-related deaths and their relationship to detention-periods varied over the phases of the conflict. As shown in Table 5, reported deaths in detention were overwhelmingly concentrated in the first and last phase of the conflict. Whereas conflict-related deaths which occurred outside of detention (for those victims who had been arbitrarily detained during the Commission's mandate period) were overwhelmingly concentrated in the first phase of the conflict: 85.3% (7651/8967) of these conflict-related deaths occurred in the first phase of the conflict, whereas 8.5% (762/8967) occurred in 1999.

Table 5: Reported Fatal Violations & Their Detention Context by phase, 1974 - 1999

	Phase 1 (1974-1983)	% Phase 1 (1974-1983)	Phase 2 (1984-1998)	% Phase 2 (1984-1998)	Phase 3 (1999)	% Phase 3 (1999)	Total
Victim detained at time of violation	216	2.4	48	6.5	114	12.1	**378**
Victim detained but unknown dates	1,172	12.8	114	15.3	28	3.0	**1,314**
Victim detained - but not at time of violation	7,651	83.8	554	74.6	762	81.2	**8,967**
Victim never detained	88	1.0	27	3.6	35	3.7	**150**
Total	**9,127**	**100**	**743**	**100**	**939**	**100**	**10,809**

Source: Database of Amnesty International Reports on East Timor

2. Non-Fatal Violations[41]

2.1 Introduction

In this section, we present our analysis of non-fatal violations which were reported to the Commission. This analysis does not include overall estimations of the total extent, pattern, and trend of non-fatal violations, as the analysis is based on a convenience sample of narrative statements collected by the Commission. However, the analysis presents the statistical patterns of non-fatal violations reported to the Commission and notes hypotheses which the data support. In addition, we compare the statistical patterns and trends observed in the Commission's data on non-fatal violations to data collected contemporaneously by Amnesty International and also data collected by the Timorese NGO, Fokupers, immediately after the UN-sponsored Popular Consultation.

2.2 Overview of Statistical findings on Non-fatal violations

This section summarizes the main findings of the commission's descriptive statistical analysis of the almost 8,000 narrative statements collected in all 13 districts.

- Non fatal violations reported to the Commission were overwhelmingly concentrated in the period of the initial invasion and occupation by the Indonesian military forces and around the time of the UN-sponsored Popular Consultation: 56.3% (33,224/60,047) of documented non-fatal violations occurred between 1975 and 1984, and 21.0% (12,634/60,047) occurred in 1999.
- In almost all districts, except for Oecusse, detention, torture and ill treatment were the mostly frequently reported violations, accounting for between 69.4% and 82.7% of the reported violation counts in districts. In Oecusse, physical integrity violations accounted

[41] The authors gratefully acknowledge comments and suggestions from Professor David Banks (Duke University) and Professor Herbert F. Spirer (University of Connecticut and Columbia University).

for 43.0% of the district's violation count. Relative to other districts, in Oecusse, property and economic violations were reported in higher proportions, comprising 30.8% (1,271/4,133) of the district's total violation count.

- The patterns of non fatal violations during the first and last phases of the conflict varied from region to region. While the initial violence around the time of the Indonesian invasion in 1975 was most intense in the Western and Central Regions, after 1976 the focus of non-fatal violations shifted to the Eastern Region.
- The documented age-sex distribution counts for arbitrary detention, torture and ill-treatment are remarkably similar, each showing that the most frequently documented victim group for these types of violations were young men of military age (between the ages of 20 and 39). Very few documented acts of detention, torture and ill treatment were experienced by female victims. In contrast, women experienced the overwhelming majority of sexually-based violations: 90.1% (769/853) of the sexually-based violations documented by the Commission involved female victims.
- The Commission's data on non-fatal violations show a general upward trend in the ratio of adults to children over time, that is the number of adult victims relative to child victims is larger in the latter part of the conflict.
- Contemporaneous reports from Amnesty International show three distinct waves of detentions of identified individuals in 1985, 1989-1993 and 1994-1999 of 402, 891 and 811 respectively. Whereas retrospective narrative statements given to the Commission suggest that the bulk of arbitrary detentions occurred in 1999 and around 1975-1984.
- The Commission's comparative analysis between its own statistical data and contemporaneous reports by Amnesty International show that although international human rights groups, such as Amnesty International, meticulously documented the human rights situation in East Timor throughout the 1980s and 1990s, there was substantial under-reporting of the overall magnitude of non-fatal violence at the time especially during the initial invasion and occupation years.
- The Commission's statistical evidence is consistent with the hypothesis that the detention practices of the Indonesian military shifted from a focus on both individual and group victims in the early occupation years of 1977-1984 to a more targeted strategy focused on individual detainees from 1985 to 1999. The Commission's statistical evidence also suggests a positive correlation between acts of torture committed against group victims and individual victims over time.
- The pattern of reported detentions and torture over time was strongly positively correlated. Over time violence became increasingly coordinated and the magnitude of reported acts of torture increased over time (between the late 1970s and mid-1980s) relative to the number of reported detentions. The Commission's statistical evidence also suggests that over time (and particularly after 1984) the practice of arbitrary detention became more targeted and was used more regularly in combination with acts of torture.
- The abuses which were most often committed during known periods of detention were torture (38.4%, 4,267/9,094), ill-treatment (33.2%, 9,094/27,998) and threats (21.3%, 634/9,094). Furthermore, torture and ill- treatment are reported much less frequently

among victims who never have been held in detention: of the torture violations documented by the commission, 16.4% (1,820/11,123) were suffered by victims who never experienced detention. The Commission's statistical evidence is consistent with the hypothesis that victims who are held in detention were at an increased level of risk of being subjected to torture or ill-treatment than individuals who were never been detained during the Commission's reference period.

- Districts which reported relatively higher proportions of torture and ill treatment, tended to also report higher proportions of abuse within detention.
- Children and older people were detained substantially less often, and when they were detained, they were subjected to proportionally lower levels of abuse.
- Data collected independently by the Commission and Amnesty International confirm that large groups of people were detained on Atauro Island in the period between 1980 and 1984, in addition to continued large-scale detentions in other parts of Timor.
- 88.7% (68,943/77,748) of non-fatal violations reported to the Commission were violations against the civilian population. However, as the pro-independence movement grew more organized and popular in the lead-up to the UN-sponsored Popular Consultation in 1999, increasing numbers of civilians with pro-independence affiliations appear to have suffered non-fatal violation.
- The overwhelming majority of non-fatal violations reported to the Commission were attributed to the Indonesian military and police: 62.2% (37,343/60,047) of documented non-fatal violations were attributed to the Indonesian military and police, 38.7% (23,253/60,047) to the Timorese auxiliaries of the Indonesian occupation force and 11.9% (7,157/60,047) to the resistance movement.[42]
- The Commission's quantitative analysis of arbitrary detentions is consistent with the hypothesis that coordination and cooperation between the Indonesian occupation force and their Timorese auxiliaries was particularly strong after the Indonesian military had secured large parts of Timor-Leste and started consolidating its occupation of the territory and then again in 1999 in the lead-up to and aftermath of the UN-sponsored Popular Consultation.
- The Commission's statistical data is consistent with the hypothesis that in 1999 the Indonesian military and police aided and abetted their Timorese auxiliaries (principally the pro-autonomy militias) in the widespread use of arbitrary detention in the lead up to and aftermath of the UN-sponsored Popular Consultation.

[42] Note that for some reported violations, perpetrator responsibility was attributed to multiple institutions. Hence the percentage share of attributed institutional perpetrator responsibility does not sum to 100%.

2.3 In-depth Descriptive Statistical Analysis of non-fatal violations

2.3.1 The Nature of the Narrative Text-based Data Sources

2.3.1.1 Geographic Coverage of the Commission's statement-taking

The Commission planned to collect testimonies from approximately 1% of the total population. The strategic plan directed that 670 statements would be collected from each of the 13 districts, regardless of that district's population size and distribution. The statement-taking process covered all 67 sub-districts[43] in each of the 13 districts of Timor-Leste. In addition to the district-level statement collection, the Commission also collected 86 <s00120> statements from East Timorese refugees in West Timor, through the Commission's partnership with a coalition of West Timor-based NGOs.[44]

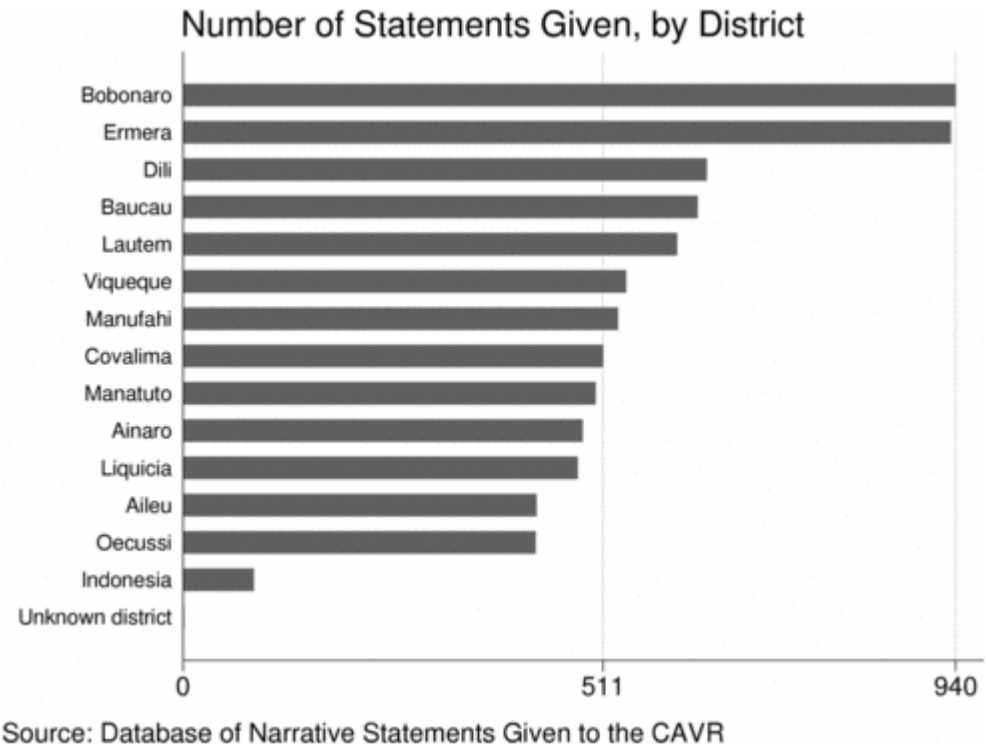
Given that the process of statement-giving was entirely voluntary, on the part of the deponent, and based on a convenience sample the distribution of statements across geographic locations was not uniform. As Figure 24 indicates, the commission collected substantially more statements from deponents in Bobonaro and Ermera than from deponents in other districts.[45]

[43] The Commission used the administrative boundary demarcations which are described in the 2001 Timor-Leste Suco Survey ADB, ETTA, UNDP and World Bank, 2001, The 2001 Survey of Sucos: Initial Analysis and Implications for Poverty Reduction, October.

[44] The Coalition of NGOs included the Center for Internally Displaced Persons Service (CIS), Truf-K, Lembaga Advokasi Kekerasan Masyarakat Sipil (Lakmas), Yabiku and Yayasan Peduli Indonesia (YPI). Staff from these NGOs collected statements from East Timorese living in Belu, Kefa, Soe and Kupang between February and August 2003.

[45] See section 2.3.1.3 below for a detailed description of the possible factors which influenced the sampling process during the Commission's statement-taking process.

Figure 24



Source: Database of Narrative Statements Given to the CAVR

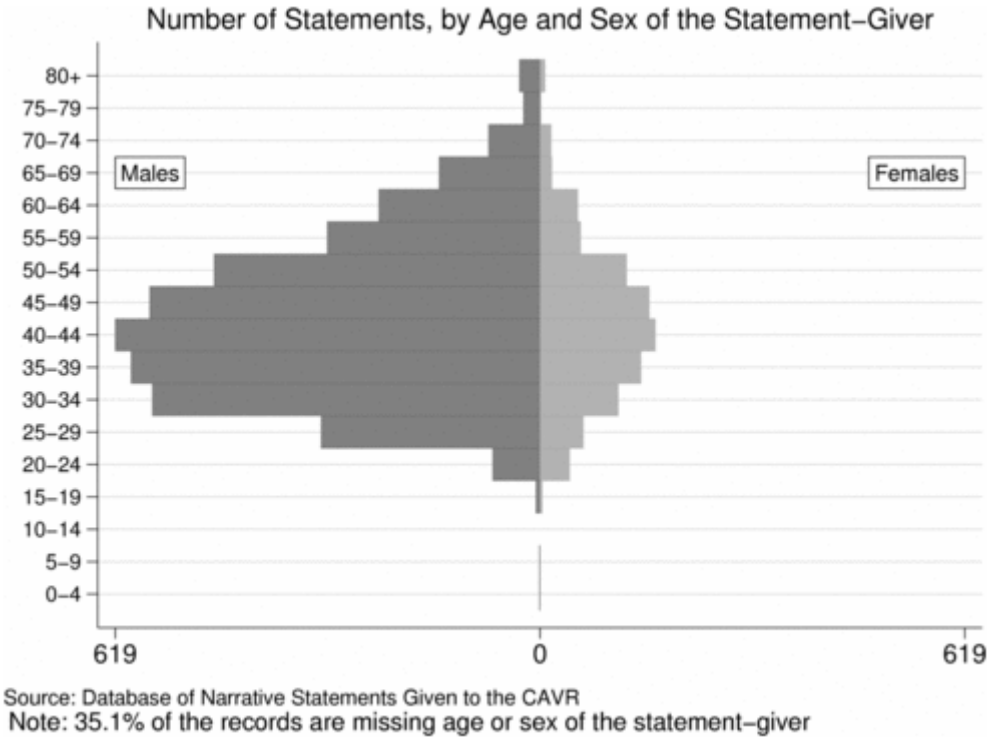
2.3.1.2 Demographic Characteristics of Deponents

Approximately 21.4% (1,642/7,669) <s00104> of all deponents in the Commission's statement-taking process were women. In some communities, women did not participate in the Commission's socialization activities as they were expected to stay at home, fewer women were organized in formal organizations with access to information regarding the Commission's work, and some women were uncertain or shy about coming forward to give testimony.[46]

The Commission received statements from adults of all ages. For both males and females, the highest number of deponents were in the 40-44 age group, as indicated in Figure 25.

[46] Pigou, Piers 'Report for CAVR's Truthseeking Division on behalf of ICTJ, (3 July - 2 August, Unpublished paper on file with CAVR.

Figure 25



Despite the substantial difference in male/female participation rates in the Commission's statement-taking process, female deponents tended to talk about violations against themselves (relative to violations against others) in roughly the same proportion as male deponents. As Table 6 shows, of all the violations reported by females, 30.6% (2,939/9,605) were violations against themselves, whereas for male deponents, 35.3% (17,438/49,382) of reported violations were against themselves.

Table 6: Count of Reported Non-Fatal Violations Cross-Tabulated by the Deponent's Sex and the Victim's Sex, 1974-1999

Deponent Sex	Deponent same as Victim	Deponent different to Victim	% (Deponent same as Victim)	% (Deponent different to Victim)	Total
Female	2,939	6,666	30.6	69.4	9,605
Male	17,438	31,944	35.3	64.7	49,382
Unknown Sex	8	979	0.8	99.2	987
Total	**20,385**	**39,589**	**34.0**	**66.0**	**59,974**

Source: Database of Narrative Statements Given to the CAVR

Women deponents often faced additional social, cultural and economic challenges relative to their male counterparts.[47] These challenges may have limited their participation in the Commission's socialization and statement-taking processes. However, the commission's statistical findings are consistent with the claim that most of the victims of killings,

[47] Ibid.

disappearances, torture and ill-treatment were young males.[48] In contrast, the overwhelming majority of sexual violations documented by the Commission were suffered by female victims, as shown in Section 2.3.5.

Statement-takers interviewed deponents in Tetum, Indonesian or any other language that could be written down. Statement-taking forms were available in Tetum and Indonesian. The 7,668 <s00101> statements received by the Commission and found to be within the CAVR mandate : 81.7% were in Tetum, 17.0% in Indonesian, 1.2% in other Timorese languages, and 0.1% in a language that was not specified. As CAVR's statement-taking forms were in Tetum and Indonesian, statements given in other languages were translated by the statement-takers and transcribed onto the official form in either Indonesian or Tetum before coding, data-entry and analysis of the narrative statements.

2.3.1.3 Potential sampling biases in the statement-taking process

As discussed in Section 2.2.1.1 the voluntary nature of the Commission's statement-taking process resulted in a degree of "self-selection." This "self-selection," in turn, introduced a number of factors which affected who was able to give a statement, such as:

(1) people who lived in remote and mountainous areas very far from where the data were being collected (such as district towns) had less chance of being in the sample than those closer to regional towns and district capitals;

(2) people who were socially active and/or physically agile were more likely to give statements than those who were sick, elderly, disabled or traumatized;

(3) people who were active in the local community and/or closely affiliated with local village, sub-district and district officials and elders were more likely to participate in the socialization process and statement taking because these local statement-collection efforts were often organized through local village structures and officials;

(4) people who died before the Commission was formed did not have an opportunity to tell their stories to the Commission; therefore, events that took place in the past, particularly in the earlier years of the Commission's reference period, tended to be less frequently reported than more recent events.

(5) people with little or no access to the media and mass communication were less likely to approach the Commission; and

(6) people from constituencies that were hostile to the Commission were less likely to make statements.

In order to address sampling biases, CAVR supplemented the statement taking process, by the collection of narrative statements from Fokupers and secondary source information from Amnesty International (see Section 2.2.2). Furthermore, to account for biases in measurement of displacement and fatal violations, the Commission developed a Retrospective Mortality Survey which collected structured information from a random probability sample of households in Timor-Leste (See Section 5.2) for a detailed presentation of the design of the sampling techniques and survey instruments which were used for the Retrospective Mortality Survey).

[48] Refer to Part 11: Acolhimento and Victim Support in Commission for Reception, Truth and Reconciliation "Chega!: The Final Report of the Commission for Reception, Truth and Reconciliation" (October, 2005).

2.3.1.4 Other non-CAVR sources

In addition to collecting its own information, the Commission collected qualitative information from local and international organizations such as Fokupers, Amnesty International and Tapol. This information was used to enrich the source material available to the Commission's researchers. The information collected from Amnesty International and Fokupers was coded using the same methods and techniques which were used for the Commission's own narrative testimony. As the volume of the Amnesty International and Fokupers data was substantially smaller than the Commission's testimony data, only focused analyses of sexual violations during 1999 and arbitrary detentions during the consolidation years were possible.

2.3.1.5 Amnesty International

The Commission received qualitative reports and Urgent Actions from the London-based human rights NGO, Amnesty International (AI). AI was the premier international human rights NGOs covering the Timorese human rights situation, using information from underground networks in East Timor and through contacts within the Timorese diaspora communities in Australia and Portugal.

The commission received 322 reports and documents from Amnesty International which were compiled between 1985 and 1999.[49]

These documents were coded and entered into the Commission's human rights violations database using the same methods and standards as were used for the 7,669 statements which were collected by the commission. The information collected from Amnesty International provides insight into the general human rights situation in East Timor as perceived by the international human rights community observed at during the Indonesian occupation, when access and information in and out of Timor was difficult. [Refer to Section 2.3.2 for more detailed discussion about the pattern of reported violations by Amnesty International.]

2.3.1.6 Fokupers

The Timorese human rights NGO, Fokupers, constructed a violations database after the referendum-related violence in 1999.[50] The Fokupers database is constructed from open-ended interviews conducted by Fokupers staff with local Timorese women. Originally the main purpose of the interviews was linked to the counseling work that Fokupers was doing. However, the objectives were extended to include documentation for investigation purposes by competent

---

[49] The Commission was unable to locate the following Amnesty International reports or Urgent Actions:
ASA 21/12/83 UA 212/83 21 September
ASA 21/16/85 Disappearances
ASA 21/44/85 Unfair Trials and Possible Torture in East Timor
ASA 21/22/87 Statement on ET by AI to the UN Special Committee on Decolonisation
ASA 21/23/87 ET: Releases of Political Prisoners
ASA 21/14/91 AI statement to UN Special Committee on Decolonisation - Appendix I and II
ASA21/24/91 East Timor: After the massacre - Appendix 1
As a result the commission's statistical analysis of violation in East Timor reported by Amnesty International does not include relevant acts and incidents covered in these reports.

[50] Fokupers was founded in 1997 and focuses on support to victims of political violence through counseling programs and other forms of assistance to women victims of violations, including ex political prisoners, war widows and wives of political prisoners. Its mandate also includes promoting women's human rights among the local population, especially East Timorese women.

legal authorities, such as the UN's Serious Crimes Unit. The narrative statements were taken in the Tetum language.

2.3.2 Overall Distribution of Reported Non-fatal Violations

The overall scale of the statement-taking process implemented by the Commission was unprecedented in scale, compared with all previous human rights documentation projects carried out in Timor-Leste.

As is shown in Table 7, the types of non-fatal violations documented by the Commission, Amnesty International and Fokupers differed significantly. This is reflective of both the differing nature of the three projects and the different social networks to which the three institutions' data collection strategies gave them access. Of all the non-fatal violations reported to the Commission, 42.3% (25,347/59,972) were detentions, 18.5% (11,123/59,972) were acts of torture and 14.1% (8,436/59,972) were acts of ill-treatment. In contrast, a substantially lower proportion of detentions (23.4% (184/788)) and tortures (7.5% (59/788)) were reported to Fokupers compared with those reported to the CAVR, although approximately similar proportions of displacements and ill-treatments were reported to Fokupers as to CAVR.

Fokupers is a women's rights NGO which also provides counseling and rehabilitation services to women. It documented a significantly larger proportion of rapes than both the CAVR and Amnesty International: 7.7% (86/1,115) of all their documented non-fatal violations were rapes. By contrast, of the violations documented in the available Amnesty reports, 59.7% (3,272/5,479) were detentions, 18% (986/5,479) were unfair trials and 11.5% (631/5,479) were acts of torture.

Table 7: Non-Fatal Violations by Datasource, 1974-1999

Violation type	CAVR Statements Database	Fokupers Database	Amnesty Database	% CAVR	% Fokupers	% Amnesty	Total
Attempted Civilian Killing	1,966	49	215	3.3	6.2	3.5	2,230
Detention	25,383	184	3,672	42.3	23.3	60.1	29,239
Torture	11,135	59	666	18.5	7.5	10.9	11,860
Rape	393	84	12	0.7	10.6	0.2	489
Sexual Slavery	98	25	1	0.2	3.2	0.0	124
Sexual Violence	221	31	43	0.4	3.9	0.7	295
Ill-Treatment	8,443	98	0	14.1	12.4	0.0	8,541
Forced Marriage	131	4	0	0.2	0.5	0.0	135
Impediments to Reproductive Rights	10	1	0	0.0	0.1	0.0	11
Unfair Trial	213	0	1,155	0.4	0.0	18.9	1,368
Destruction of Homes	2,231	51	25	3.7	6.5	0.4	2,307
Destruction of Livestock	409	6	0	0.7	0.8	0.0	415
Extortion	2,095	44	19	3.5	5.6	0.3	2,158
Threats	2,987	136	64	5.0	17.2	1.1	3,187
Forced Recruitment	2,157	10	7	3.6	1.3	0.1	2,174
Forced Labor	2,175	9	0	3.6	1.1	0.0	2,184
Other	0	0	227	0.0	0.0	3.7	227
Total	**60,047**	**791**	**6,106**	**100.0**	**100.0**	**100.0**	**66,944**

Source: Database of CAVR, Fokupers and Amnesty International Statements and Reports

The broad relative distributions of victims per violation for the different violation types was fairly similar for the CAVR, Fokupers and Amnesty projects, as shown in Tables 8, 9 and 10. The CAVR project tended to document slightly more violations per victim than the Fokupers and Amnesty projects. On average, 2.36 violations per victim were reported to CAVR, compared with 2.01 and 1.53 respectively for Fokupers and Amnesty. This difference reflects the different character of the different projects. CAVR documented violations across the entire mandate

period, including the initial invasion years, while Amnesty's work was concentrated mostly on the consolidation years of the occupation and was compiled during the conflict when communication between Timor and the rest of the world was severely restricted. The Fokupers project focused almost exclusively on the third phase of the conflict around the time of the UN-supervised Popular Consultation. Fokupers relied exclusively on female deponents and was focused on documenting sexual violations.

Table 8: Count of Reported Non-Fatal Violations by Violation-Type, 1974-1999

Violation type	Count of violations	Percent of violations	Count of victims	Percent of victims	Violations per victim
Detention	25,347	42.3	17,169	67.4	1.48
Torture	11,123	18.5	8,508	33.4	1.31
Ill-Treatment	8,436	14.1	6,872	27.0	1.23
Property/Economic Violations	4,735	7.9	3,851	15.1	1.23
Other	4,339	7.2	4,030	15.8	1.08
Threats	2,982	5.0	2,653	10.4	1.12
Forced Recruitment	2,157	3.6	1,988	7.8	1.09
Sexually-Based Violations	853	1.4	657	2.6	1.30
Total	**59,972**		**25,460**		**2.36**

Note: the victim proportions sum to more than 100% because the same victim may suffer more than one violation type

Source: Database of Narrative Statements Given to the CAVR

Table 9: Count of Reported Non-Fatal Violations by Violation-Type 1974-1999

Violation type	Count of violations	Percent of violations	Count of victims	Percent of victims	Violations per victim
Detention	184	23.4	168	42.7	1.10
Sexually-Based Violations	145	18.4	104	26.5	1.39
Threats	136	17.3	119	30.3	1.14
Property/Economic Violations	99	12.6	86	21.9	1.15
Ill-Treatment	98	12.4	92	23.4	1.07
Torture	59	7.5	59	15.0	1.00
Other	57	7.2	54	13.7	1.06
Forced Recruitment	10	1.3	10	2.5	1.00
Total	788		393		2.01

Note: the victim proportions sum to more than 100% because the same victim may suffer more than one violation type

Source: Database of Narrative Statements Given to Fokupers

Table 10: Count of Reported Non-Fatal Violations by Violation-Type 1974-1999

Violation type	Count of violations	Percent of violations	Count of victims	Percent of victims	Violations per victim
Detention	3,272	59.7	3,073	86.0	1.06
Other	1,412	25.8	1,383	38.7	1.02
Torture	631	11.5	600	16.8	1.05
Threats	61	1.1	59	1.7	1.03
Sexually-Based Violations	55	1.0	43	1.2	1.28
Property/Economic Violations	44	0.8	43	1.2	1.02
Forced Recruitment	4	0.1	4	0.1	1.00
Total	5,479		3,572		1.53

Note: the victim proportions sum to more than 100% because the same victim may suffer more than one violation type

Source: Database of Amnesty International Reports on East Timor

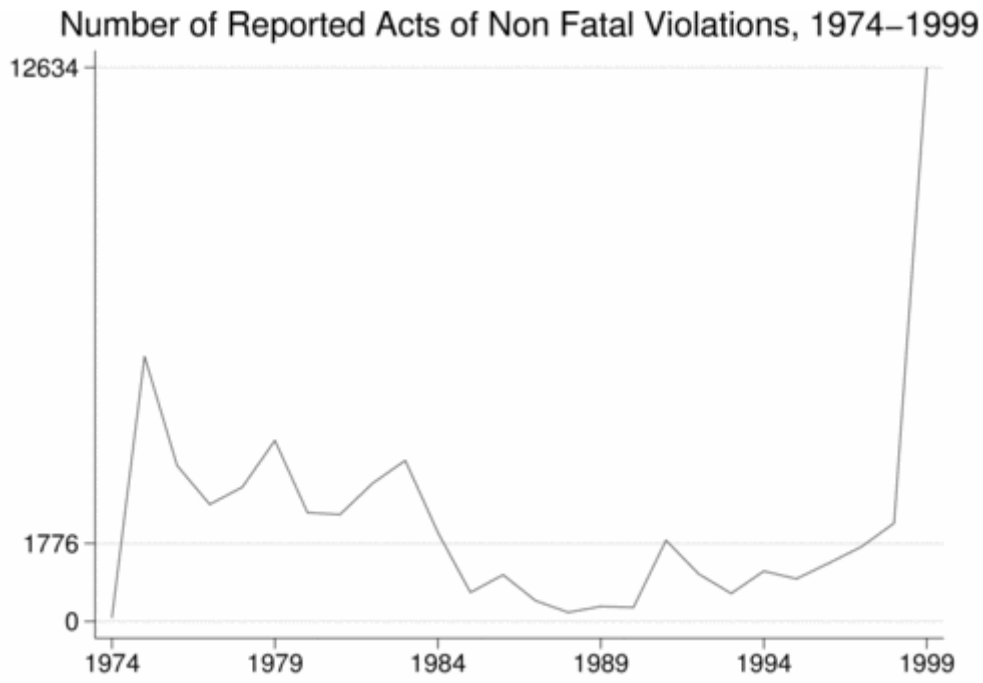
2.3.3 The three phases of large-scale violence in East Timor

The Commission defines three phases of conflict during April 1974 – September 1999. The first phase includes the initial Indonesian invasion and occupation of East Timor, spanning 1975 to 1984. The second phase is the consolidation and normalization of the occupation, from 1985 to

1998. The third phase of conflict includes the first three quarters of 1999, the period surrounding the UN-sponsored Popular Consultation process.

As can be seen in Figure 26, there were high levels of non-fatal violations during the initial invasion and occupation. During the second phase in general there were relatively lower levels of non-fatal violations and a concentration of detentions, tortures and killings around the time of the 1991 Santa Cruz massacre. The second phase of normalization included a new wave of targeted detentions and physical abuse of suspected members and collaborators with the resistance movement. Finally the last phase of the conflict, which includes the lead-up to the Popular Consultation and also the period between the Popular Consultation and the deployment of the multinational Interfet (International Force in East Timor), produced two distinct waves of killing, displacement and looting. This final phase was characterized by large-scale violations concentrated in a short period of time overwhelmingly carried out by "pro-autonomy militias" supported, trained, armed and directed by the Indonesian military.

Figure 26



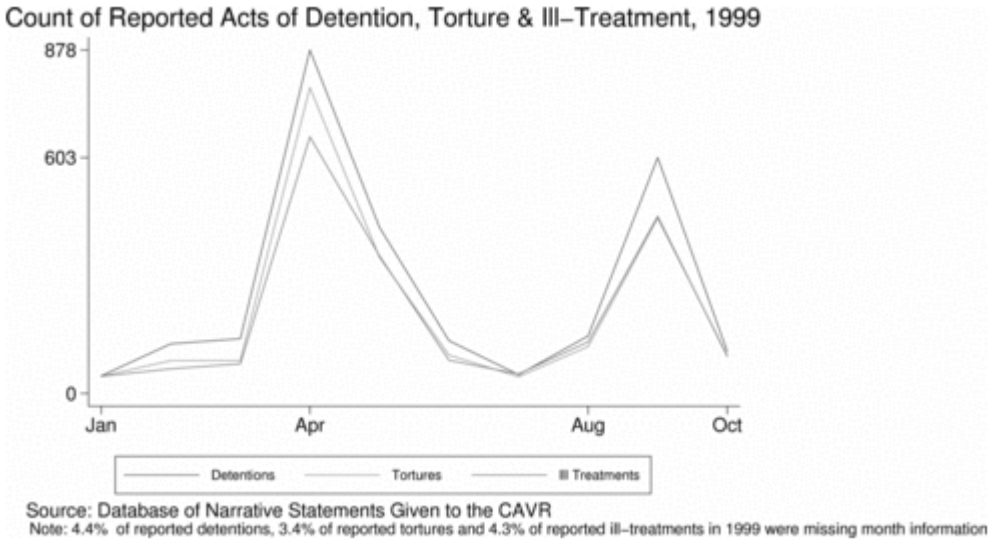
Source: Database of Narrative Statements Given to the CAVR

The pattern of massive non-fatal violations during the initial invasion and occupation years, followed by relatively low-level violence during the consolidation and normalization years and then an increase of violence in 1999 is also mirrored in the pattern of fatal violations over time, as discussed in Sections 1.4.1 and 1.4.2 above.

During 1999, reported violations were overwhelmingly concentrated in April and September. As Figure 27 shows, the reported pattern of detention, ill-treatment and torture are positively correlated over time, that is, when any one of the violations increases, the others also tend to increase, and vice-versa. All three violation types have reported peaks in April with a slightly smaller peak being reported in September, although both peaks are of a similar magnitude. In 1999 reported violence was concentrated into two main bursts before and after the process

leading to the UN-sponsored Popular Consultation. This pattern is consistent with the hypothesis that physical intimidation was used in a coordinated fashion to intimidate the Timorese population in the lead-up to the arrival of the UN mission, UNAMET, that was authorized to conduct the referendum and as retribution in the immediate aftermath of the ballot.

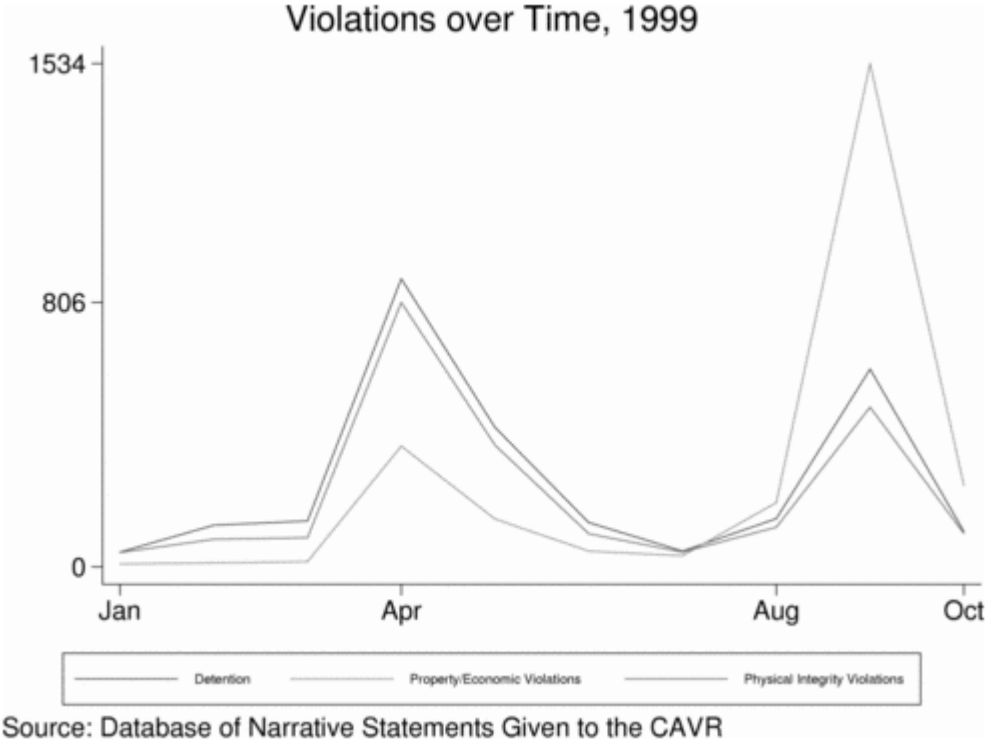
Figure 27



Source: Database of Narrative Statements Given to the CAVR
Note: 4.4% of reported detentions, 3.4% of reported tortures and 4.3% of reported ill–treatments in 1999 were missing month information

There was a substantial shift in practice by the Indonesian-backed militias and the Indonesian military before and after the Popular Consultation. As Figure 28 shows, the most frequently used form of repression used prior to the ballot were physical integrity violations (such as detention, torture and ill treatment). Immediately after the referendum, looting and forms of property and economic violations were used most frequently. After the ballot, physical integrity violations occurred at a slightly lower level compared with the pre-ballot period, but they are overshadowed by property violations. This pattern appears to be consistent with the hypothesis that the Indonesian military and militias switched from using violations of physical integrity before the ballot to pressure the population to vote for autonomy to retributive acts after the ballot result consisting of large-scale looting and property destruction coupled with retributive acts of physical violence.

Figure 28



Source: Database of Narrative Statements Given to the CAVR

The Commission's empirical data on non-fatal violation patterns over time in 1999 are consistent with the hypothesis that violence was coordinated in 1999.

2.3.4 Variations in reported non-fatal abuses across space

The Commission's narrative data cannot be used to assess directly the differences in the magnitude of violations between regions and districts. Data based on convenience samples are representative only of the total extent of violence from region to region in so far as the deponents whose statements were taken are representative of their local population and were selected in proportion to the violence suffered in each district. As described in Section 5.1.0 below, the narrative information collected by the Commission, Fokupers and Amnesty International all are subject to a number of biases. Consequently, patterns of non-fatal violations across space are presented in this section in order to gain insight into the social processes of data collection by the Commission, Fokupers and Amnesty International, and to assess whether the reported patterns across space are consistent with relevant qualitative analysis and argument.

Table 11: Count of Reported Non-Fatal Violations by Geographic Location, 1974-1999

Violation	Lautem	Viqueque	Baucau	Manatuto	Manufahi	Aileu	Ermera	Liquiça
Detention	2,696	2,412	2,366	1,414	2,237	1,520	2,010	1,074
Physical Integrity Violations	909	905	976	695	1,094	658	1,281	662
Property/ Economic Violations	80	610	289	103	189	123	578	103
Other	1,319	1,570	1,205	893	1,954	1,013	2,112	858
Total	**5,004**	**5,497**	**4,836**	**3,105**	**5,474**	**3,314**	**5,981**	**2,697**

Table 11 (continued)

Violation	Dili	Ainaro	Covalima	Oecussi	Bobonaro	Indonesia	Total
Detention	4,658	1,611	935	501	1,718	195	25,347
Physical Integrity Violations	1,367	743	641	460	1,035	90	11,516
Property/ Economic Violations	229	232	227	1271	505	196	4,735
Other	2,135	1,141	1,267	1,166	1,494	247	1,8374
Total	**8,389**	**3,727**	**3,070**	**3,398**	**4,752**	**728**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

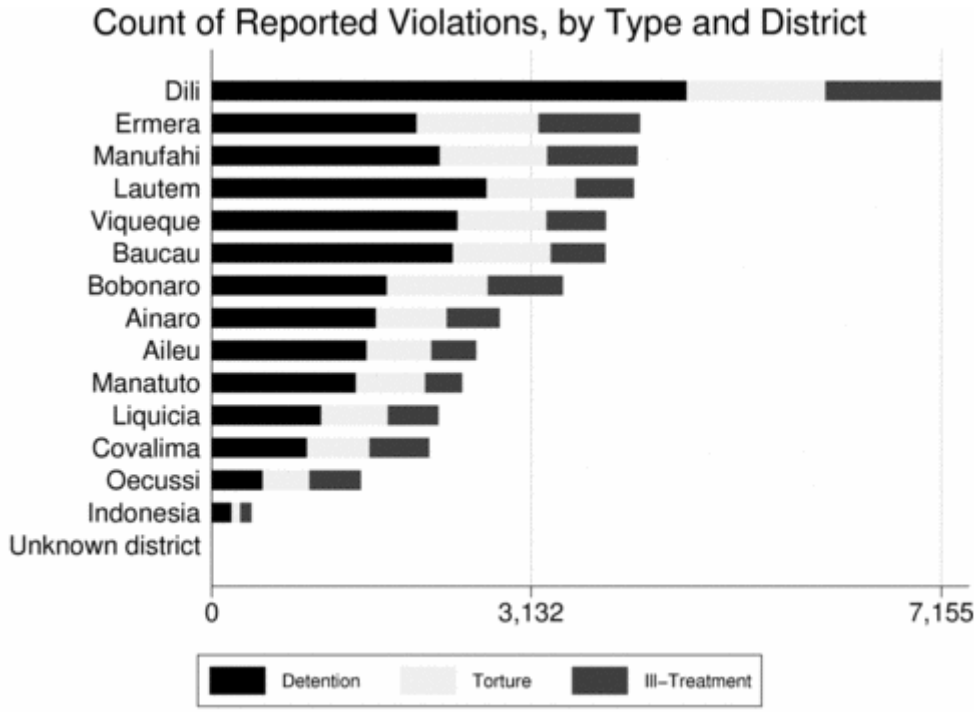
Table 11 shows the counts of each violation by district reported during the Commission's statement-taking process. Dili has a significantly higher number of reported violation counts than any other district, comprising 14.0% (8,389/59,972) all violations in the country. The districts with other relatively high violation counts are Ermera, Manufahi, Viqueque and Lautém. In almost all districts, except for Oecusse, detention, torture and ill treatment were the mostly frequently reported violations, accounting for between 69.4% and 82.7% of the reported violation counts in districts. In Oecusse, physical integrity violations accounted for 43.0% of the district's violation count. Relative to other districts, in Oecusse, property and economic violations were reported in higher proportions, comprising 30.8% (1,271/4,133)of the district's total violation count.[51] Property and economic violations in other districts were reported significantly less frequently on average comprising 7.4% (3,464/56,574) of reported violations.

Although physical integrity violations reported to the Commission constituted 61.5% (36,911/60,047) of all documented non-fatal violations, detention, torture and ill-treatment were not documented in the same proportions in each district, as shown in Figure 29 . In particular

[51] 96.8% (1,230/1,271) of these property violations in Oecusse were reported to have occurred in 1999. Furthermore 94.0% (3,194/3,398) of reported violations in Oecusse occurred in 1999. Hence it appears that, unlike other districts, the violence in Oecusse was almost exclusively in 1999.

Dili has a higher proportion of documented detentions relative to the number of its documented acts of ill-treatment and detention, whereas Bonbonaro, Ainaro, Aileu, Manatuto, Liquiça and Covalima have proportionally lower numbers of documented detention compared with their respective proportions of ill-treatment and torture. The Commission's data are consistent with the hypothesis that policies and practices of detention and physical abuse varied across regions. In particular our empirical analysis found that while detentions were used more often in Dili, ill-treatment and torture were used less frequently there relative to the rest of the country.

Figure 29

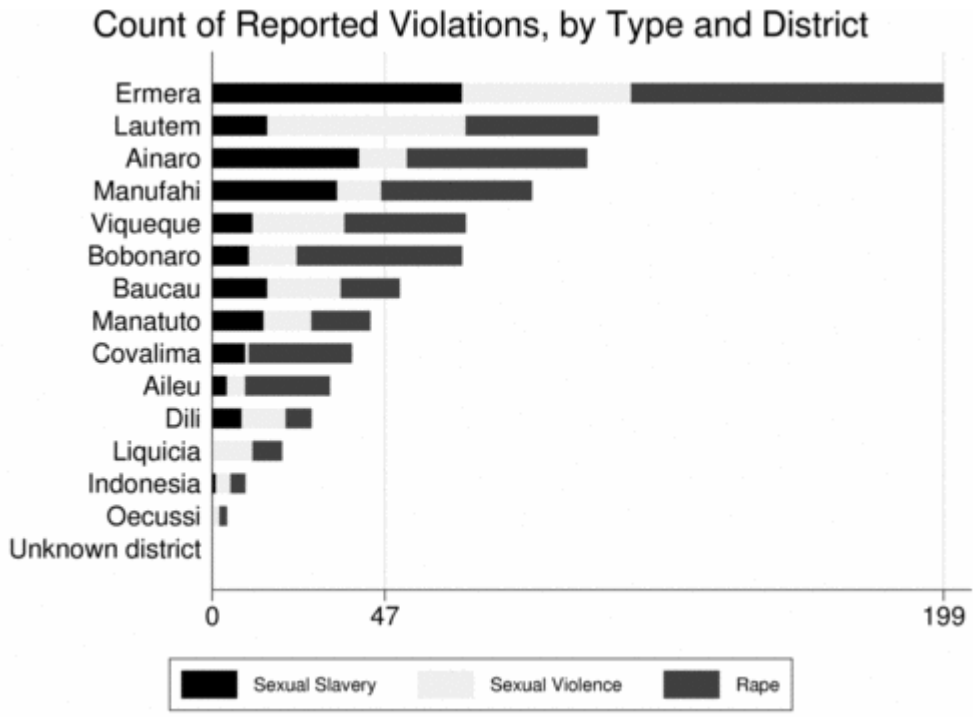


Source: Database of Narrative Statements Given to the CAVR

Sexually-based violations documented by the Commission comprised 1.4% (853/59,972) of all reported violations. However, in Ermera, Ainaro and Lautem, Commission researchers found a relatively higher proportion of sexual violations?3.3% (199/5,981), 2.7% (102/3,727) and 2.1% (105/5,004) of the total reported violations respectively. Sexual violations were reported less frequently in Dili and Oecusse representing 0.3% (27/8,389) and 0.1% (4/3,398) of the total respectively.

The types of documented sexually-based abuses varied across districts, as shown in Figure 30. Across Timor, of all the sexual violations documented by the Commission, rape accounted for 46.1% (393/853), other sexual violence 27.1% (231/853) and sexual slavery 26.8% (229/853). Rapes accounted for a higher proportion of sexual violations in Aileu and Bobonaro than the national average: 71.9% (23/32) and 66.2% (45/68), respectively. Whereas sexual slavery accounted for a higher proportion of sexually- based violations in Manufahi and Ainaro than the national average: 39.1% (34/87) and 39.2% (40/102), respectively. Similarly, other sexual violence accounted for 57.9% (11/19) and 51.4% (54/105) of all documented sexual violations in Liquiça and Lautem, respectively.

Figure 30



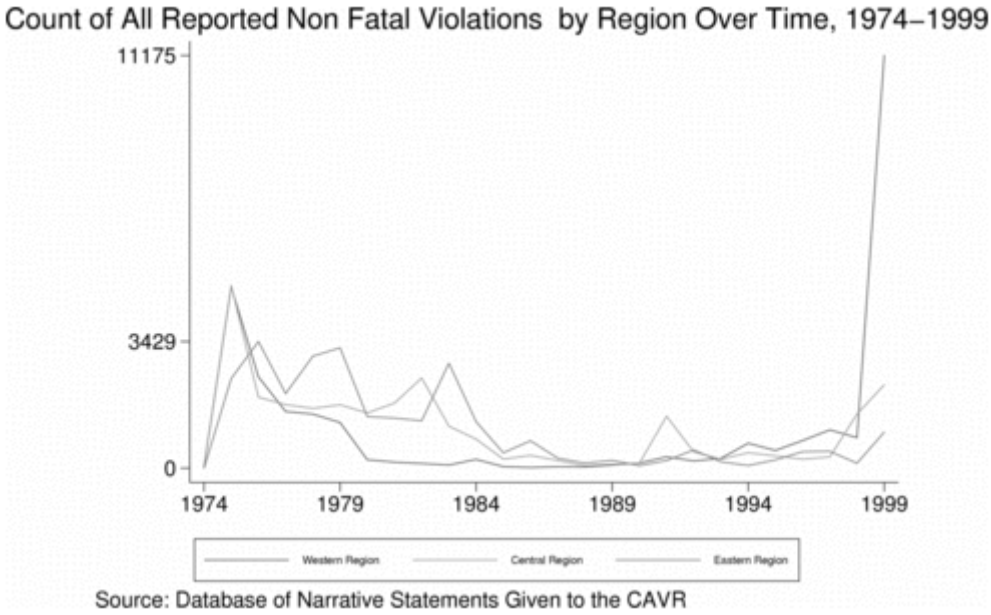
Source: Database of Narrative Statements Given to the CAVR

2.3.5 Non-Fatal Violations over Time and Space

Broadly speaking violence in Timor-Leste occurred in distinct phases, as discussed in Section 2.3.3. However, the patterns of non fatal violations during the first and last phases of the conflict varied from region to region as shown in Figure 31. In particular, violence associated with the initial Indonesian invasion and party conflict in 1975 was more intense in the Western and Central Regions relative to the Eastern Region. However, as the occupation continued, reported non-fatal abuses in the Western Region decreased from its initial high levels in 1975 to a relatively low level by 1980, whereas in the Central Region violence also decreased after the initial invasion period to a level of intensity of about half that experienced in 1975. In the Eastern Region the level of documented violence in 1975 was only about as half as much in absolute terms as that reported in the Western and Central Regions. However, across Timor-Leste, throughout the late 1970s and early 1980s violence continued at around the same levels as was experienced in 1976 without any substantial decrease until 1984. While the initial violence around the time of the Indonesian invasion in 1975 was most intense in the Western and Central Regions, after 1976 the focus of non-fatal violations shifted to the Eastern Region.

Apart from the Santa Cruz massacre and its aftermath in 1991 in Dili, reported violence during the "consolidation years" from 1984 through to 1998 took the form of sporadic low-level violence in all three regions. During the final phase of the conflict in 1999, 75.1% (9,494/12,634) of reported non-fatal violations occurred in the Western Region. The Commission's data are consistent with the claim that populations close to the West Timorese border and in Oecusse in 1999 were subjected to higher levels of violence as the pro-autonomy militias and Indonesian military withdrew towards West Timor. [Cross-reference to Section in CAVR Report on 1999 violence]

Figure 31



2.3.6 Age-Sex Victim Demographics of Reported Non Fatal Violations

2.3.6.1 Reporting Levels of Age and Sex Information for Victims

The Commission examined several hypotheses that might establish whether or not victims were targeted on the grounds of age and sex. This section describes the notably different demographic age-sex patterns for reported victims of the main non-fatal violations. This analysis includes only victims whose age and sex at the time of the violation was known.

Of the 60,047 non-fatal violations reported to the Commission, 34,047 contained exact age information of identified victims (63.4%) of non-fatal violations. Unfortunately we have no way of knowing what the distribution of unknown ages is for these data. Hence it is not possible to assess how representative the age distributions of victims with known ages are of the age distribution of all reported victims.

The Commission considered a child to be any person under the age of 18 years old. This definition conforms with the definition set out in the United Nations Convention on the Rights of the Child.[52] The majority, 89.8% (30,574/34,047), of non-fatal violations documented by the Commission, where the victim's age was known were perpetrated against adults. 10.2% (3,473/34,047) of violations documented by the Commission were suffered by child victims. The Commission documented 26,000 non-fatal violations against victims whose exact age was unknown.

Of the 60,047 non-fatal violations documented by the Commission, 99.4% (59,715/60,047) were against victims whose sex was known. Of these violations 14.0% (8,355/59,715) were committed against females and 86.0% (51,360/59, 715) were committed against males. 25,476 victims

[52] Article 1, Convention on the Rights of the Child, UN General Assembly Document A/RES/44/25 (12 December 1989) with Annex.

(including those whose sex was not known) experienced these documented violations: of whom 15.7% (4,002/25,476) were females and 83.6% (21,308/25,476) were male.

The Commission documented 3,473 violations against children, of which 3,451 violations have known information about the sex of the child. Of these violations, 27.5% (950/3,451) were against girl victims and 72.5% (2501/3,451) were against boy victims. There were 22 children whose sex was either unknown or not reported by the witness. Of the 30,446 documented violations against adults where the sex of the victims is known, 12.7% (3,870/30,446) were females and 87.3% were males. Thus, the proportion of documented violations against female children is greater than the proportion of documented violations against female adults. Hence, both adult and child victims tend to be males. Relative to males, female victims tended to be younger.

2.3.6.2 Victim Analysis by Sex

The types of violations perpetrated against males and females are substantially different. In Table 12, it is clear that females suffer the overwhelming majority of sexual violations: for every one sexually-based violation against a male the Commission documented ten violations against females. Whereas for every act of torture and forced recruitment against a female victim, the Commission documented around 12 or 13 acts of tortures and forced recruitment against males. Other types of violation, such as threats, property and economic violations, ill-treatment and detention were documented in an average proportion of about 5.8 male victims for each female victim.

Table 12: Count of Reported Non-Fatal Violations by Violation Type and Sex of Victim, 1974-1999

Violation Type	Violations against Males	Violations Against Females	Violations against Unknown Sex	% Female Victims	Proportion Male to Female Victims	Total
Property/Economic Violations	3,792	908	35	19.2	4.2	4,735
Physical Integrity Violations	10,205	1,247	64	10.8	8.2	11,516
	15,602	2,678	94	14.6	5.8	18,374
Detention	21,687	3,521	139	13.9	6.2	25,347
Total	**51,286**	**8,354**	**332**	**13.9**	**6.1**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

In Table 13, we analyze the pattern of victims by sex over time. There is significant variation in the ratio of male to female victims across the different years of the conflict: the male-to-female ratio varies from a low of 2.7 in 1981 to a high of 43.1 in 1991. Higher than average male-to-female ratios are recorded in 1975, 1987, and 1999. These findings are consistent with the hypotheses that substantial numbers of females were transported to Atauro in 1981 and that nature of the post-Santa Cruz crackdown by the Indonesian military forces was largely directed against males.

Table 13: Count of Reported Non-Fatal Violations by Year and Sex of Victim, 1974-1999

Year	Violations against Males	Violations Against Females	Violations against Unknown Sex	% Female Victims	Proportion Male to Female Victims	Total
1974	94	0	0	0.0		94
1975	5,526	489	29	8.1	11.3	6,044
1976	2,986	554	13	15.6	5.4	3,553
1977	2,287	352	30	13.2	6.5	2,669
1978	2,516	523	23	17.1	4.8	3,062
1979	3,471	633	23	15.3	5.5	4,127
1980	2,071	389	20	15.7	5.3	2,480
1981	1,768	664	7	27.2	2.7	2,439
1982	2,440	708	4	22.5	3.4	3,152
1983	2,949	679	41	18.5	4.3	3,669
1984	1,788	239	2	11.8	7.5	2,029
1985	532	128	0	19.4	4.2	660
1986	856	204	2	19.2	4.2	1,062
1987	430	41	0	8.7	10.5	471
1988	189	16	0	7.8	11.8	205
1989	311	20	11	5.8	15.6	342
1990	302	7	8	2.2	43.1	317
1991	1,658	185	6	10.0	9.0	1,849
1992	1,009	54	20	5.0	18.7	1,083
1993	551	84	2	13.2	6.6	637
1994	974	158	14	13.8	6.2	1,146
1995	927	38	6	3.9	24.4	971
1996	1,170	154	6	11.6	7.6	1,330
1997	1,580	123	1	7.2	12.8	1,704
1998	2,075	164	4	7.3	12.7	2,243
1999	10,826	1,748	60	13.8	6.2	12,634
Total	**51,286**	**8,354**	**322**		**6.1**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

Although a district breakdown shows less variation across space in male-to-female victim ratios than over time, higher than average ratios were documented in Liquiça (11.4), Oecusse (9.4), Aileu (8.3) and Dili (8.2), whereas Lautem (3.7), Ainaro (4.5) and Ermera (4.5) recorded lower than average male-to-female ratios. These findings are shown in Table 14.

Table 14: Count of Non-Fatal Violations by Geographic Location and Sex of Victim, 1974-1999

District	Violations against Males	Violations Against Females	Violations against Unknown Sex	% Female Victims	Proportion Male to Female Victims	Total
Lautem	3,915	1069	20	21.4	3.7	5,004
Viqueque	4,766	685	46	12.5	7.0	5,497
Baucau	4,049	745	42	15.4	5.4	4,836
Manatuto	2,566	511	28	16.5	5.0	3,105
Manufahi	4,829	621	24	11.3	7.8	5,474
Aileu	2,934	352	28	10.6	8.3	3,314
Ermera	4,872	1,086	23	18.2	4.5	5,981
Liquiça	2,475	217	5	8.0	11.4	2,697
Dili	7,441	906	42	10.8	8.2	8,389
Ainaro	3,034	681	12	18.3	4.5	3,727
Covalima	2,624	426	20	13.9	6.2	3,070
Oecussi	3,057	325	16	9.6	9.4	3,398
Bobonaro	4,097	631	24	13.3	6.5	4,752
Indonesia	627	99	2	13.6	6.3	728
Total	**51,286**	**8,354**	**322**	**13.9**	**6.1**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

2.3.6.3 Victim Analysis by Age

In this section we examine victims' age by violation types, time and space.

The counts of specific violations by adults and children are given in Table 15. For almost all the violation types documented by the Commission, for every one violation suffered by a child, approximately 7-10 violations against adults are documented. However, for sexual violations the proportion of adult to child victims was substantially lower than other violation types: for every sexually-based violation documented by the Commission against a child, 3.4 violations against adults were documented. Hence, the adult-to-child ratio was about 2.5 times lower for sexual violations than for other non-fatal violations.

Table 15: Count of Reported Non-Fatal Violations by Violation Type and Age of Victim, 1974-1999

Violation Type	Violations against Adults	Violations Against Children	Violations against Unknown Age	% Child Victims	Proportion Adult to Child Victims	Total
Property/ Economic Violations	2,882	323	1,530	6.8	8.9	4,735
Physical Integrity Violations	6,255	639	4,622	5.5	9.8	11,516
	9,543	1,088	7,743	5.9	8.8	18,374
Detention	11,849	1,427	12,071	5.6	8.3	25,347
Total	**30,529**	**3,477**	**25,966**	**5.8**	**8.8**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

Table 16 shows the counts of violations against adults and children by year. The Commission's data on non-fatal violations show a general upward trend in the ratio of adults to children over time, that is the number of adult victims relative to child victims is larger in the latter part of the conflict. However, since there is a substantially larger amount of "age-missingness" for victims in the earlier part of the conflict, it is difficult to make comparisons between the adult-to-child victim ratio in the early and late periods of the conflict.

Table 16: Count of Reported Non-Fatal Violations by Year of Violation and Age of Victim, 1974-1999

Year	Violations against Adults	Violations Against Children	Violations against Unknown Age	% Child Victims	Proportion Adult to Child Victims	Total
1974	5	0	89	0.0		94
1975	2,616	402	3,026	6.7	6.5	6,044
1976	1,582	293	1,678	8.2	5.4	3,553
1977	1,087	185	1,397	6.9	5.9	2,669
1978	1,302	226	1,534	7.4	5.8	3,062
1979	2,015	260	1,852	6.3	7.8	4,127
1980	1,155	173	1,152	7.0	6.7	2,480
1981	1,174	293	972	12.0	4.0	2,439
1982	1,381	199	1,572	6.3	6.9	3,152
1983	1,653	200	1,816	5.5	8.3	3,669
1984	1,017	78	934	3.8	13.0	2,029
1985	307	40	313	6.1	7.7	660
1986	527	93	442	8.8	5.7	1,062
1987	194	27	250	5.7	7.2	471
1988	121	6	78	2.9	20.2	205
1989	234	17	91	5.0	13.8	342
1990	196	45	76	14.2	4.4	317
1991	1,099	62	688	3.4	17.7	1,849
1992	509	28	546	2.6	18.2	1,083
1993	308	47	282	7.4	6.6	637
1994	568	53	525	4.6	10.7	1,146
1995	457	43	471	4.4	10.6	971
1996	680	39	611	2.9	17.4	1,330
1997	1,073	88	543	5.2	12.2	1,704
1998	1,070	55	1,118	2.5	19.5	2,243
1999	8,199	525	3,910	4.2	15.6	12,634
Total	**30,529**	**3,477**	**25,966**	**5.8**	**8.8**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

On average the Commission documented 8.8 adult victims for every one child victim. However, there is variation in the adult-to-child victim ratio between districts, as is shown in Table 17. In Bobonaro a relatively high number of child victims were documented, as is shown by the reported adult-to-child victim ratio of 4.8., whereas in Covalima, Indonesia and Oecusse a noticeably higher than average adult-to-child victim ratios was documented.[53]

[53] On average, the adult-to-child victim ratio documented by the Commission in Covlima was 17.3, in Indonesia was 15.3 and in Oecusse was 14.1.

Table 17: Count of Reported Non-Fatal Violations by Geographic Location of Violation and Age of Victim, 1974-1999

District	Violations against Adults	Violations Against Children	Violations against Unknown Age	% Child Victims	Proportion Adult to Child Victims	Total
Lautem	2,498	347	2,159	6.9	7.2	5,004
Viqueque	2,119	226	3,152	4.1	9.4	5,497
Baucau	2,597	314	1,925	6.5	8.3	4,836
Manatuto	1,347	191	1,567	6.2	7.1	3,105
Manufahi	2,505	246	2,723	4.5	10.2	5,474
Aileu	1,488	166	1,660	5.0	9.0	3,314
Ermera	3,107	315	2,559	5.3	9.9	5,981
Liquiça	1,810	166	721	6.2	10.9	2,697
Dili	4,089	448	3,852	5.3	9.1	8,389
Ainaro	1,811	232	1,684	6.2	7.8	3,727
Covalima	1,942	112	1,016	3.6	17.3	3,070
Oecussi	2,278	161	959	4.7	14.1	3,398
Bobonaro	2,510	525	1,717	11.0	4.8	4,752
Indonesia	428	28	272	3.8	15.3	728
Total	**30,529**	**3,477**	**25,966**	**5.8**	**8.8**	**59,972**

Source: Database of Narrative Statements Given to the CAVR

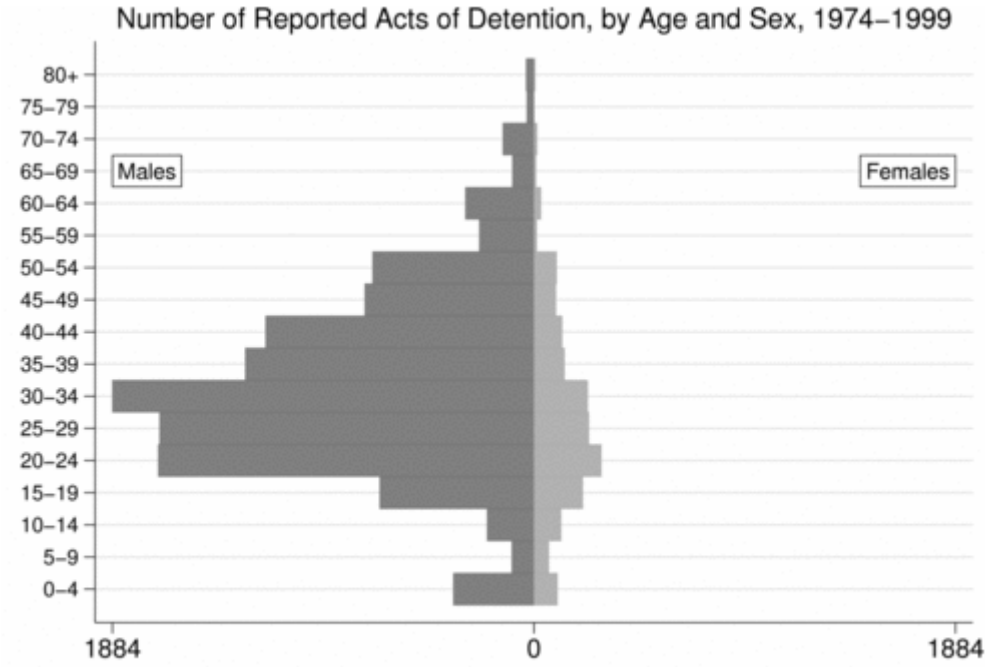
2.3.6.4 Victim Analysis by Age & Sex

This section describes the distribution of victims by both age and sex. The analysis is presented as both counts and in terms of population-based rates of each violation's occurrence. The population-based rates are calculated using the 1990 Indonesian Population Census.[54]

Figures 32, 33, and 34 present counts of documented age-sex violations for detention, torture and ill-treatment. The documented age-sex distribution counts for these three violation types are remarkably similar, each showing that the most frequently documented victim group for these types of violations were young men of military age. Very few documented acts of detention, torture and ill treatment were experienced by female victims.

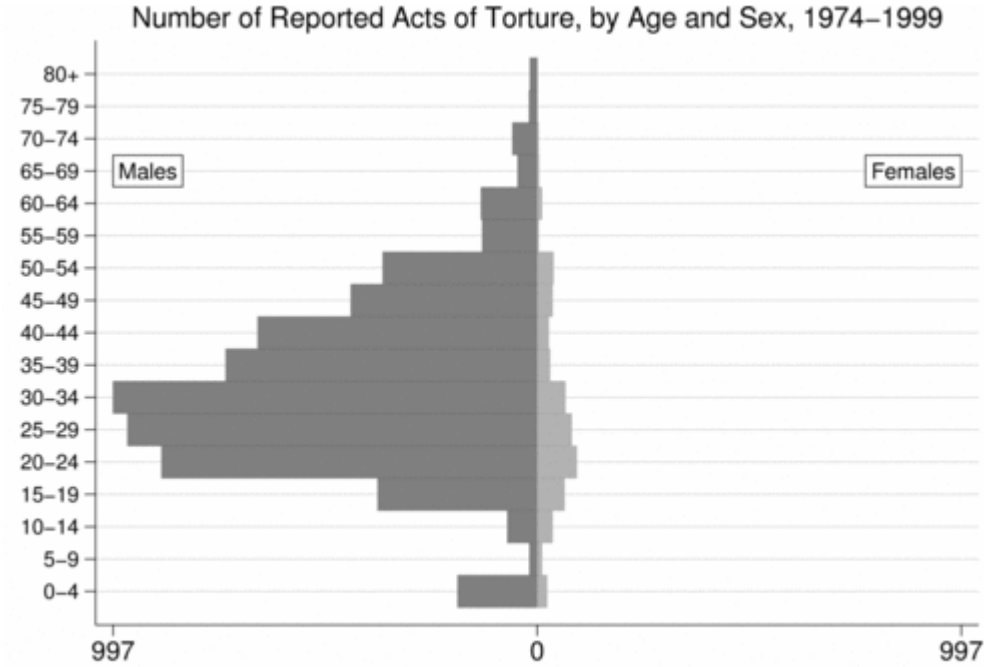
[54] BPS (Biro Pusat Statistik, Central Statistical Bureau) 1990

Figure 32



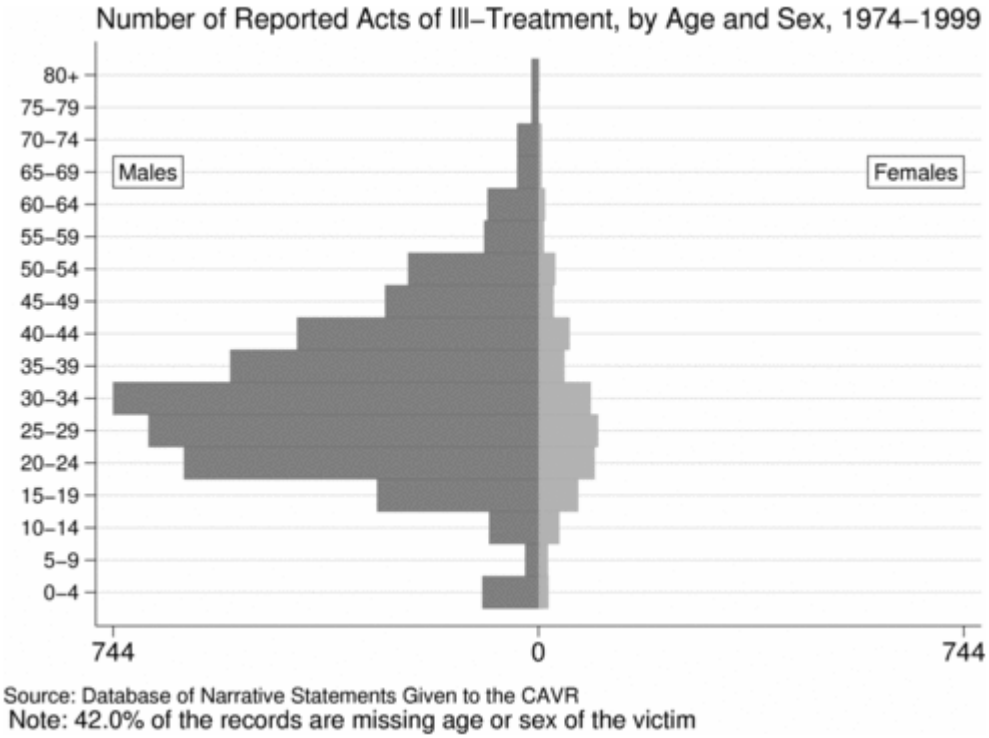
Source: Database of Narrative Statements Given to the CAVR
Note: 47.8% of the records are missing age or sex of the victim

Figure 33



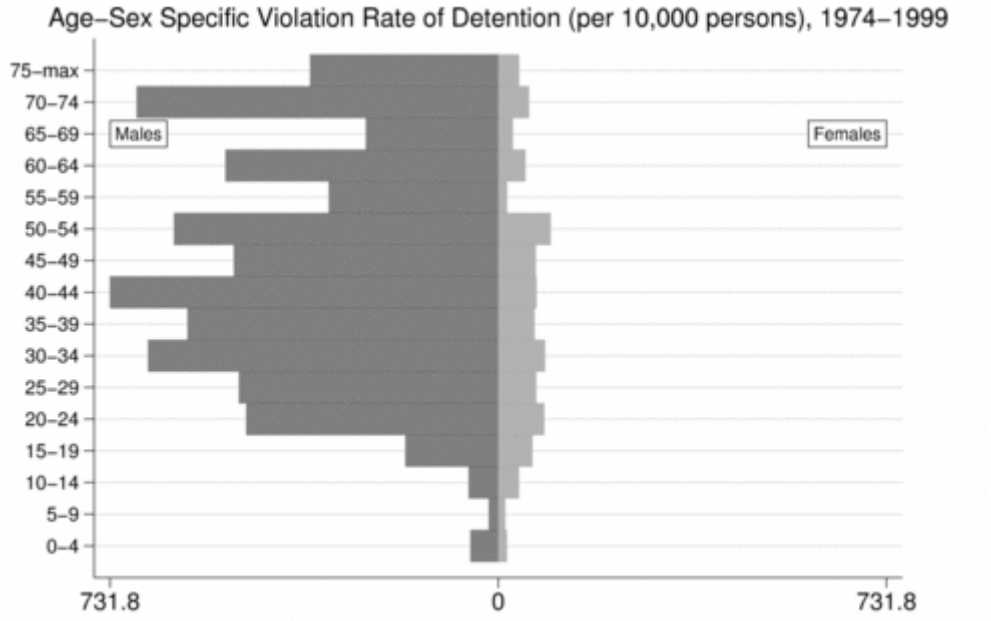
Source: Database of Narrative Statements Given to the CAVR
Note: 40.6% of the records are missing age or sex of the victim

Figure 34



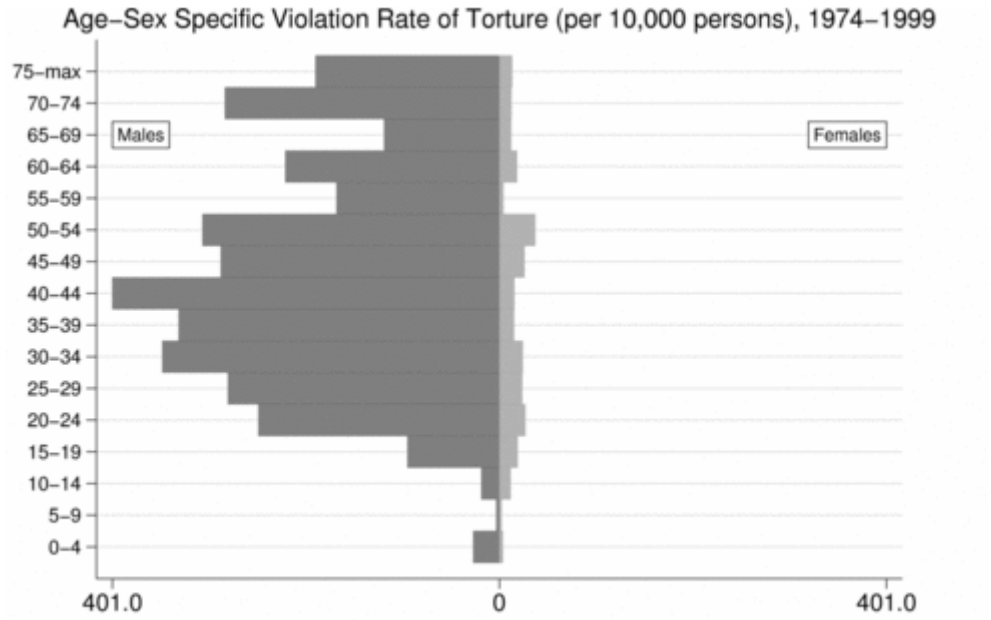
When we move from simple violation counts to population-based violation rates, we see that relative to the overall Timorese population middle-age males experienced the highest rates of these forms of violence. Furthermore, old males above the age of 70 experienced these forms of violence at a similar rate to middle-aged males. These patterns are shown in Figures 35, 36 and 37.

Figure 35



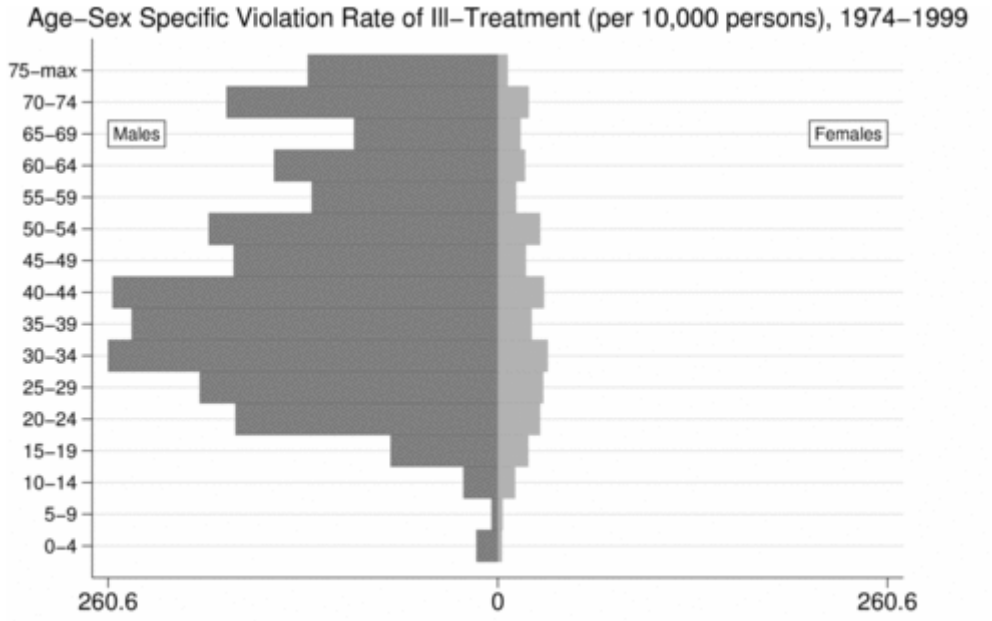
Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 47.8% of the violation records are missing age or sex of the victim

Figure 36



Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 40.6% of the violation records are missing age or sex of the victim

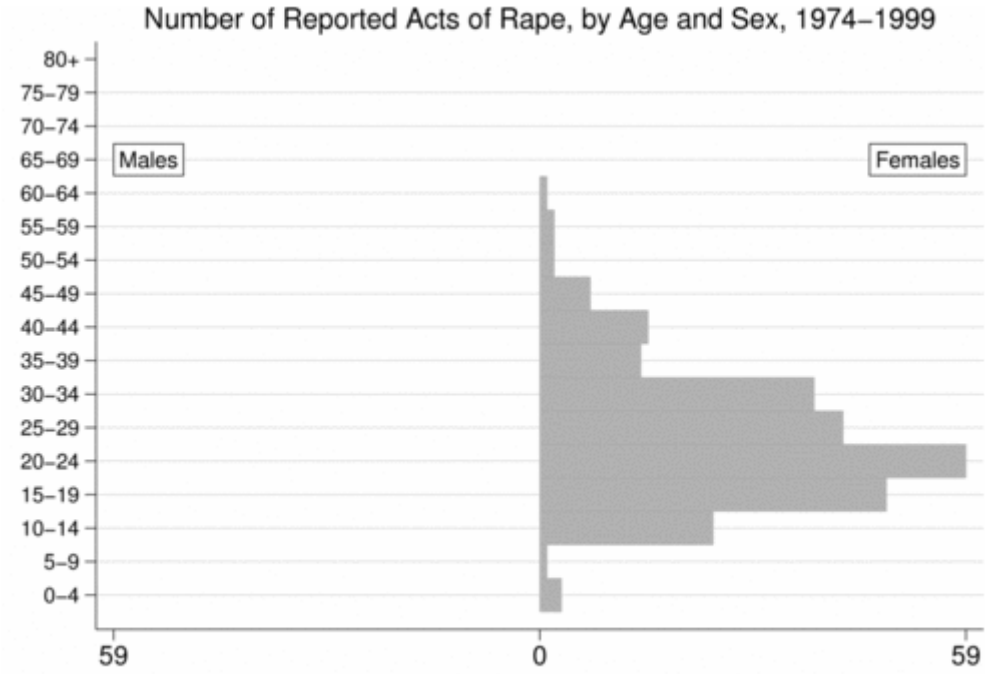
Figure 37



Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 42.0% of the violation records are missing age or sex of the victim

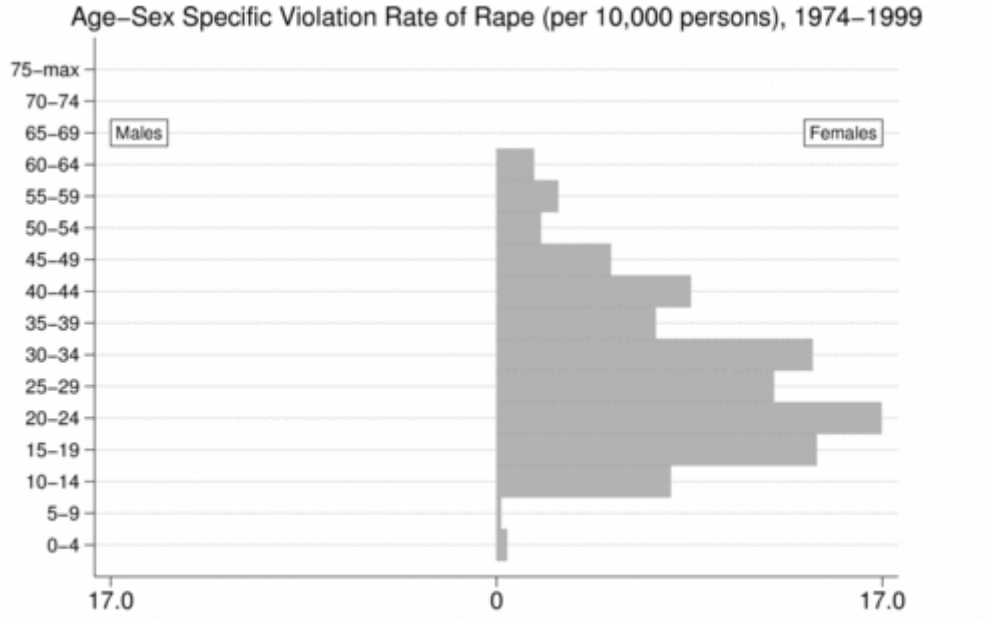
The age-sex distributions of victims of sexual violations documented by the Commission are substantially different to those for physical integrity violations. This can be seen in Figures 38 and 39. Furthermore, there are notable differences in the age-sex distribution of victims for the different forms of sexual violations. The Commission documented rapes of women in all age categories under 65 years old. However, the highest frequency of documented rape and highest population-based rates of rape were for young women of reproductive age. 15-24 year old women appear to have been the sub-population at most risk of rape.

Figure 38



Source: Database of Narrative Statements Given to the CAVR
Note: 34.9% of the records are missing age or sex of the victim

Figure 39

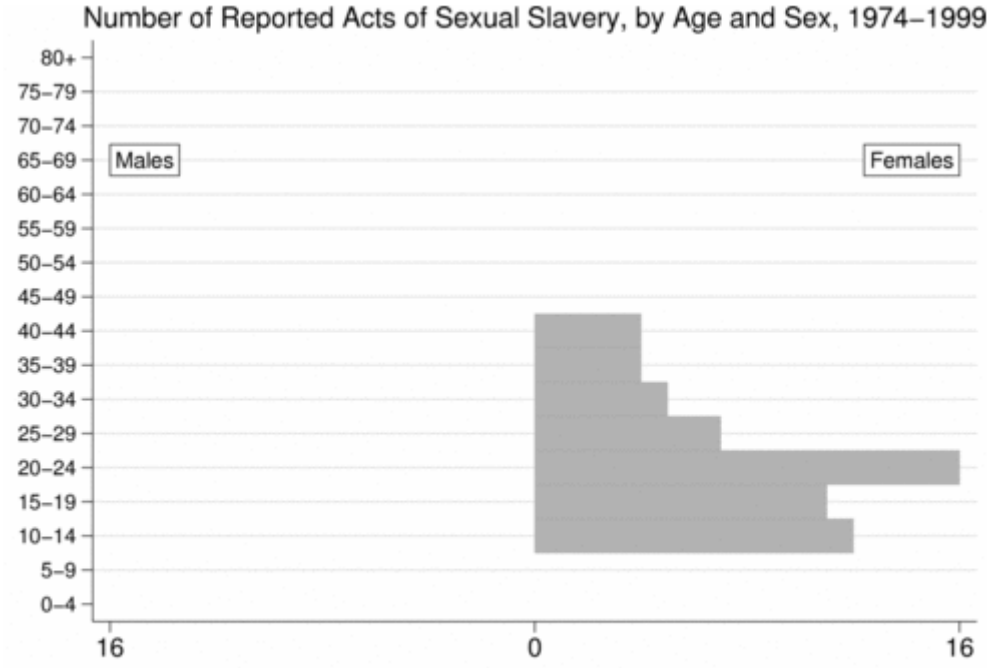


Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 34.9% of the violation records are missing age or sex of the victim

By contrast only women between the ages of 10 and 44 were among the documented victims of sexual slavery. Of these victims women between 20 and 24 years old experienced both the

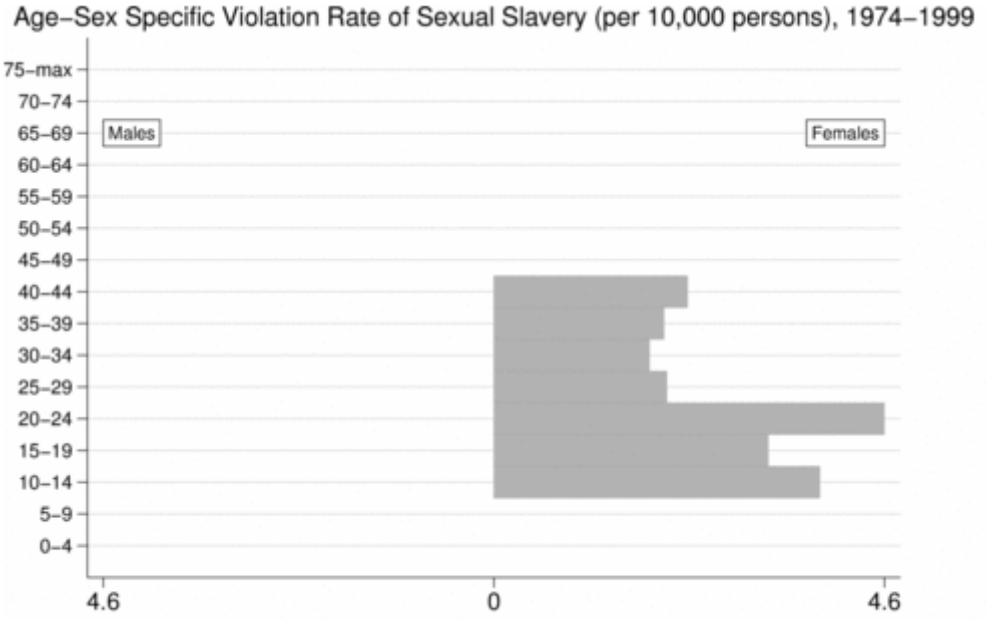
highest counts and highest rates of sexual slavery. As was the case for rape, no cases of sexual slavery of men were documented by the Commission.

Figure 40



Source: Database of Narrative Statements Given to the CAVR
Note: 39.8% of the records are missing age or sex of the victim

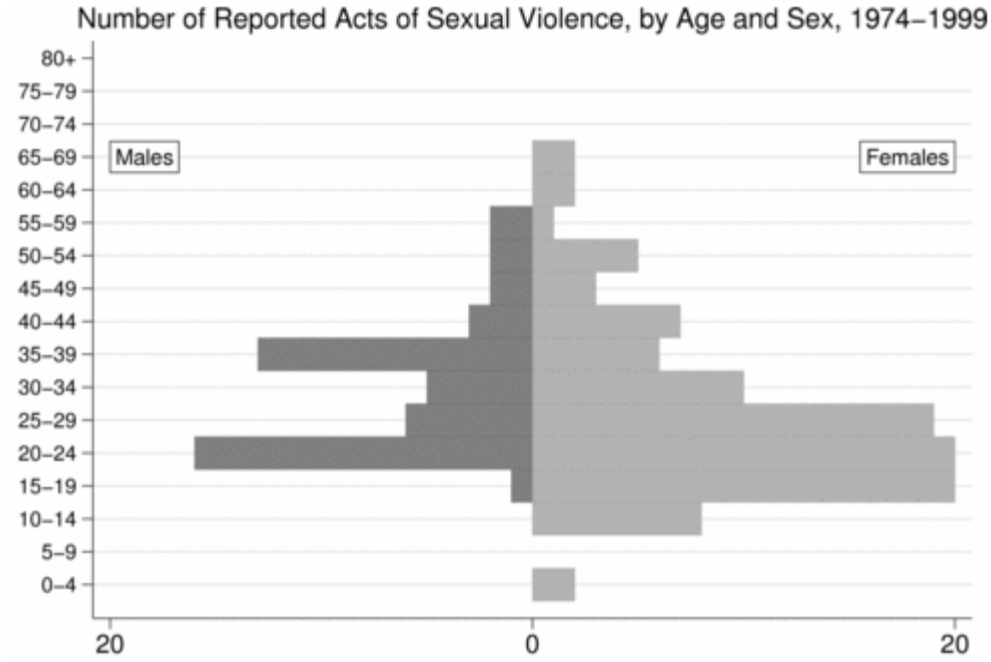
Figure 41



Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 39.8% of the violation records are missing age or sex of the victim

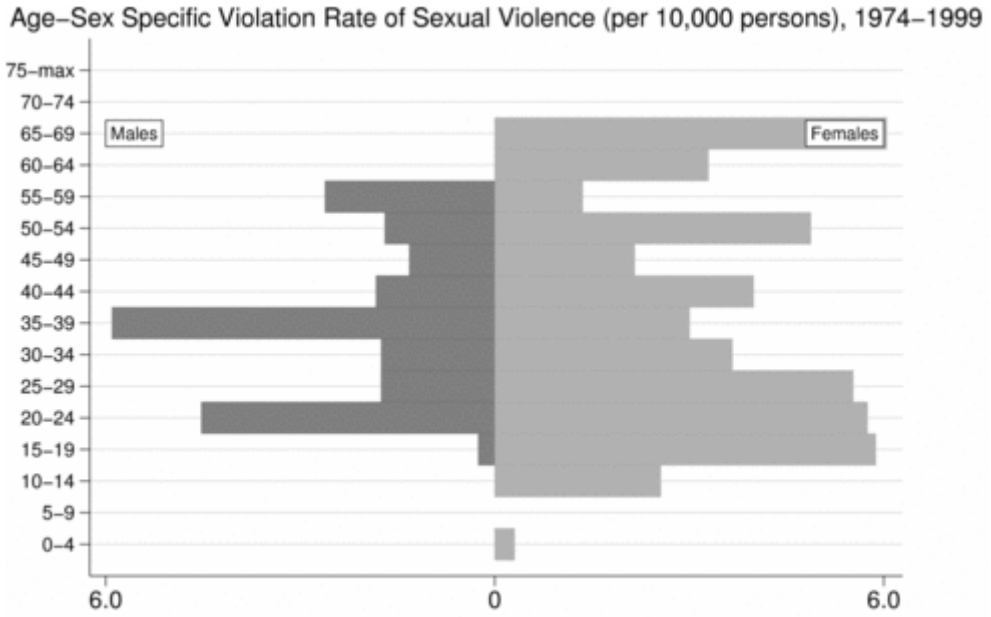
However, the Commission documented cases of other sexual violence against both men and women. This form of violence was most commonly directed against men in the 20-24 and 35-39 age groups and women between the ages of 15 and 29 years old.

Figure 42



Source: Database of Narrative Statements Given to the CAVR
Note: 29.9% of the records are missing age or sex of the victim

Figure 43



Sources: (i) Database of Narrative Statements Given to the CAVR, (ii) 1990 Population Census, Statistics Indonesia
Note: 29.9% of the violation records are missing age or sex of the victim

Hence, the Commission's quantitative analysis suggests that young women experienced the overwhelming majority of sexual violations. Furthermore, rape and sexual slavery were exclusively reported to have been suffered by women.

2.3.7 Comparison of Retrospective and Contemporaneous Human Rights Monitoring

This section compares extent and pattern of non-fatal violations reported by the Commission with the extent and pattern reported by Amnesty International. It shows how isolated Timor-Leste was from the international community and the paucity of information and limited extent to which knowledge of violations in Timor were known during the early and harshest periods of the conflict.

As discussed in Section 2.3.1, access to the territory during the Indonesian occupation was extremely limited, especially for international human rights groups such as Amnesty International (AI). As a result the geographic coverage of contemporaneous reports by AI is significantly different to that reported to the Commission and Fokupers in their retrospective statement-taking processes.

35.6% (1,953/5,479) of reported non-fatal violations documented by Amnesty International did not contain information about the location where the respective violation took place, as is shown in Table 18. This appears to be consistent with the limited information flow out of Timor-Leste during the occupation (particularly from remote, mountainous villages and sub-districts). Furthermore, the nature of contemporaneous reporting to the international community was more focused on reporting the nature of the human rights situation in Timor rather than describing the differing conditions in different parts of Timor. However, 32.3% (1,770/5,479) of the non-fatal violations reported by AI occurred in Dili; a higher proportion than was reported the retrospective projects conducted by CAVR and Fokupers in which violations in Dili accounted for 14.0% (8,389/59,972) and 4.6% (36/788).

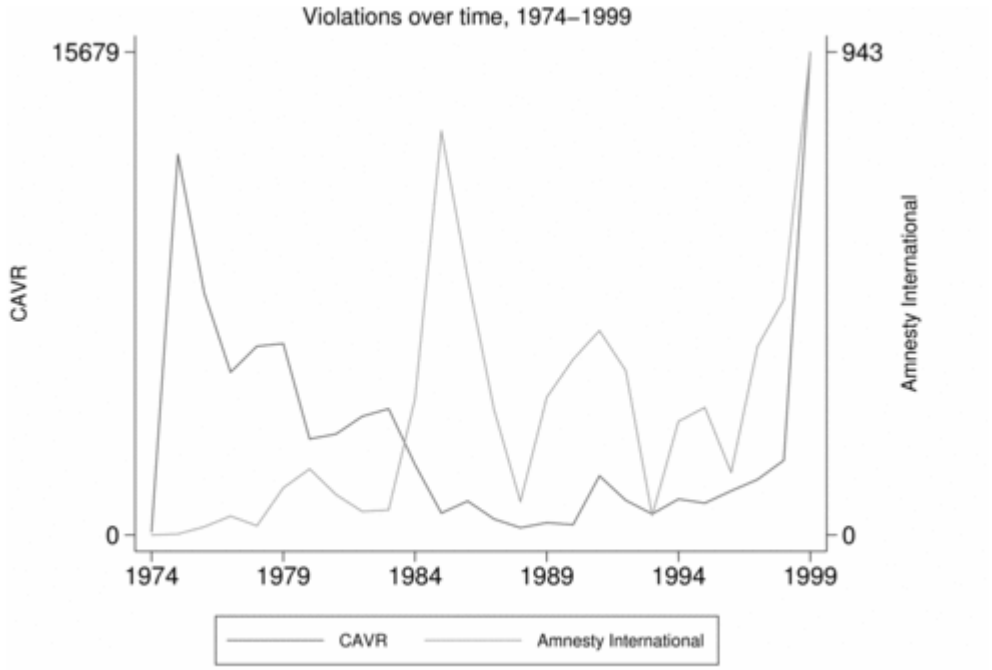
Table 18: Count of Reported Violations by Geographic Location of Violation and Datasource, 1974-1999

District	CAVR Statements Database	Amnesty Database	Fokupers Database	% CAVR	% Amnesty	% Fokupers
Lautem	5,004	6	111	8.3	2.0	0.8
Viqueque	5,497	6	114	9.2	2.1	0.8
Baucau	4,836	1	210	8.1	3.8	0.1
Manatuto	3,105	27	22	5.2	0.4	3.4
Manufahi	5,474	63	60	9.1	1.1	8.0
Aileu	3,314	18	57	5.5	1.0	2.3
Ermera	5,981	113	25	10.0	0.5	14.3
Liquiça	2,697	88	246	4.5	4.5	11.2
Dili	8,389	36	1,770	14.0	32.3	4.6
Ainaro	3,727	84	54	6.2	1.0	10.7
Covalima	3,070	90	65	5.1	1.2	11.4
Oecussi	3,398	10	19	5.7	0.3	1.3
Bobonaro	4,752	190	184	7.9	3.4	24.1
Indonesia	728	52	589	1.2	10.8	6.6
Unknown district	0	4	1,953	0.0	35.6	0.5
Total	**59,972**	**788**	**5,479**	**100.0**	**100.0**	**100.0**

Source: Database of CAVR, Fokupers and Amnesty International Statements and Reports

As indicated in Figure 44, contemporaneous reports from AI show three distinct waves of detentions of identified individuals in 1985, 1989-1993 and 1994-1999 of 402, 891 and 811 respectively. Whereas retrospective reports given to the Commission suggest that the bulk of arbitrary detentions occurred in 1999 and around 1975-1984. Furthermore, these reports suggest that at least 2,779 detentions occurred in 1999 and at least 16,509 between 1975 and 1984. These comparisons are evidence of the difficult nature of documenting human rights abuses in Timor-Leste during the Indonesian military occupation. In particular, Figure 44 shows that although international human rights groups, such as Amnesty International, meticulously documented the human rights situation in East Timor throughout the 1980s and 1990s, there was substantial under-reporting of the overall magnitude of non-fatal violence at the time.

Figure 44



Source: Database of CAVR, Fokupers and Amnesty International Statements and Reports

2.3.8 The Nature of Abuses against Individuals and Groups

Some non-fatal violations were reported to the Commission as being perpetrated against a lone individual, whereas some other violations were reported as being perpetrated against multiple individuals at the same time. Figures 45, 46 and 47 show the distribution of violations by victim group size for torture, detention and ill-treatment violations.[55]

The nature of the abuses committed against single individuals tended to be distinct from the nature of those committed against groups. As Figure 45 and Figure 47 show, arbitrary detention and ill-treatment were more commonly reported as having been perpetrated against individuals or groups of 50 or more – with fewer people being detained or suffering ill-treatment in groups of 2-49 persons. Nearly all reported acts of torture were committed against individual victims one-by-one. This empirical finding appears to be consistent with the hypothesis that the use of torture as a form of oppression was used in a more targeted fashion.

[55] As is the case with reported violations against individual victims, violations against victims in groups can be reported by more than one deponent. The Commission matched group victim records to identify duplicate reports of the same violation and victim in multiple statements. The methods used for matching are described in the Statistical Appendix in Section 5.4.

Figure 45

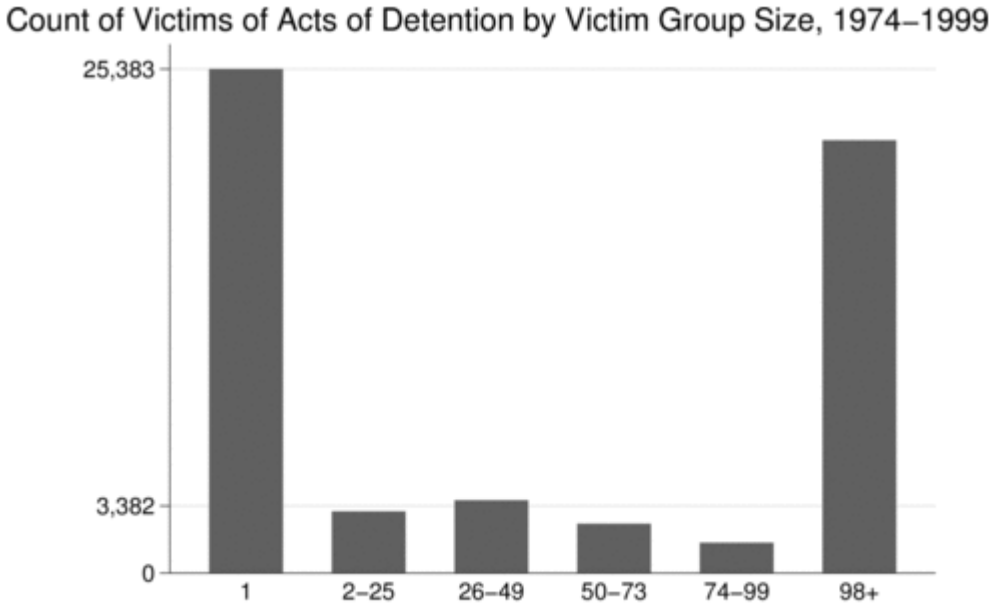


Figure 46

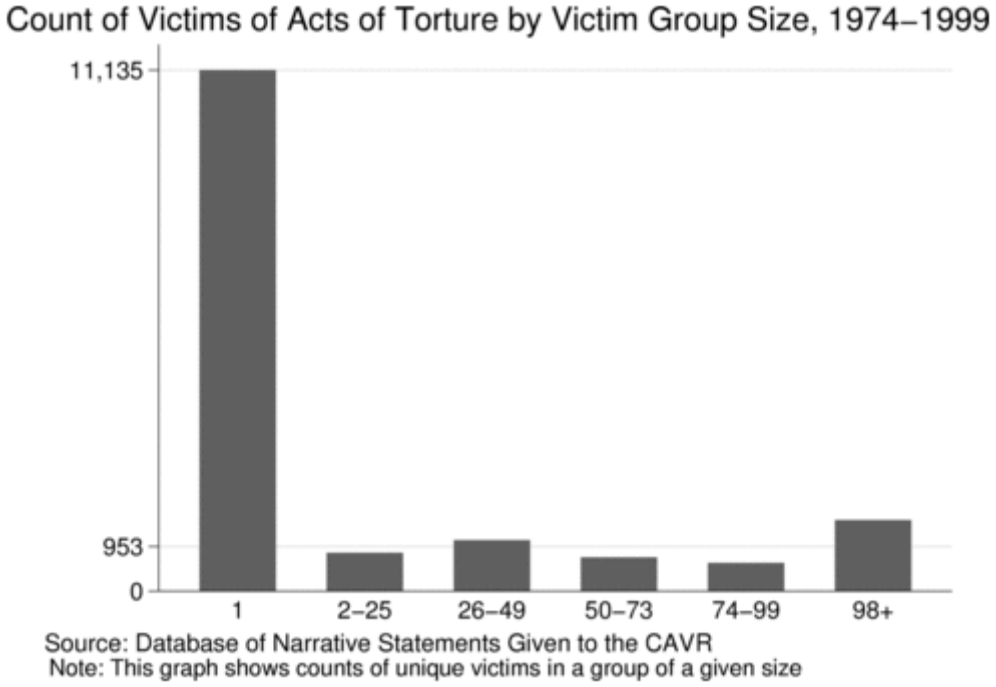


Figure 47

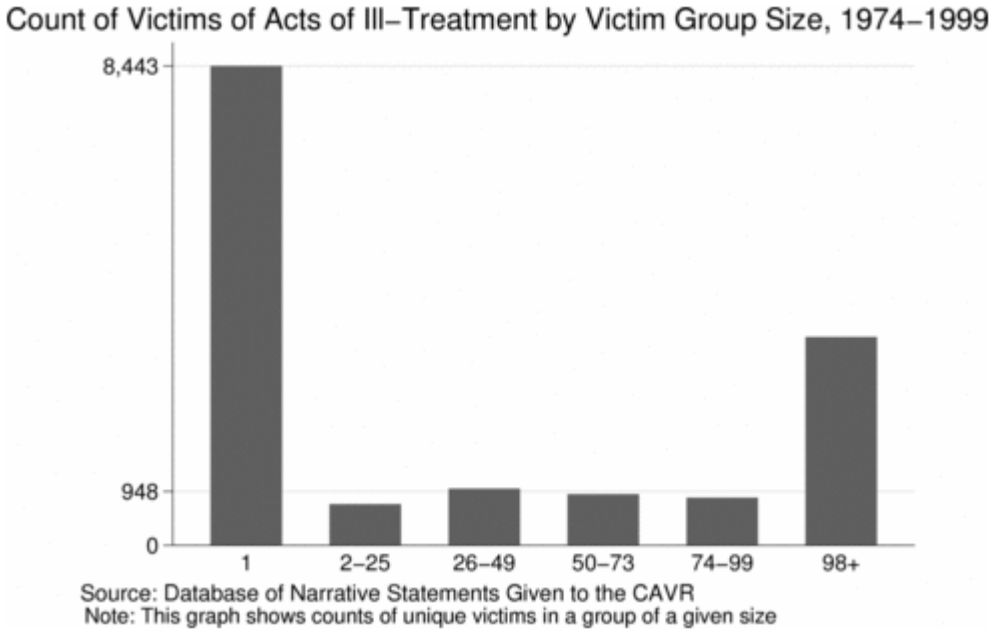


Figure 48 shows that in statements given to the Commission, the detention of individual victims and group victims are positively correlated. When reported detentions against individuals increase, so do reported detentions against group victims.[56] Furthermore, there is a substantial difference between the extent of reported detentions of individual and group victims. Between 1974 and 1984 reported detentions against group victims are almost always substantially higher than detentions of individual victims.[57] Two possible explanations for this pattern are the following.

[56] The correlation coefficient for the two series is 0.74.

[57] The only year in this period where reported detentions against group victims were fewer than reported detentions against individual victims was 1983.

Figure 48

Source: Database of Narrative Statements Given to the CAVR

i. The pattern could reflect the increasingly targeted nature of Indonesian military's detention practices over the course of the occupation and its increased ability in the 1980s and 1990s to target individuals who were contributing the Resistance movement's activities).

ii. Alternatively, deponents to the Commission's statement-taking process may have had more difficulty specifically identifying individual detainees detained in the earlier occupation years relative to detainees in later years. Consequently, deponents reporting on the earlier period may more frequently describe earlier detentions as anonymous groups.

However, it seems unlikely that the pattern among detentions is an artifact of respondent recall because none of the other violation types (namely detentions, ill-treatments, threats and property violations) exhibits any evidence of such a recall bias – as is shown in Figures 49, 50, 51 and 52.[58] Hence the statistical evidence on detentions documented by the Commission is consistent with the hypothesis that the detention practices of the Indonesian military shifted from a focus on both individual and group victims in the early occupation years of 1977-1984 to a more targeted strategy focused on individual detainees from 1985 to 1999.

The Commission's statistical evidence also suggests a positive correlation between acts of torture committed against group victims and individual victims over time.[59] Furthermore, as shown in Figure 49, peaks in reported acts of torture against group victims occurred in 1975, 1982 and

[58] See also Figures 88, 89, 90 and 91 in Section 5.4.4.
[59] The correlation coefficient for the series is 0.69.

1999. Hence, the Commission's data suggest that the bulk of mass violence against groups was heavily concentrated in time.

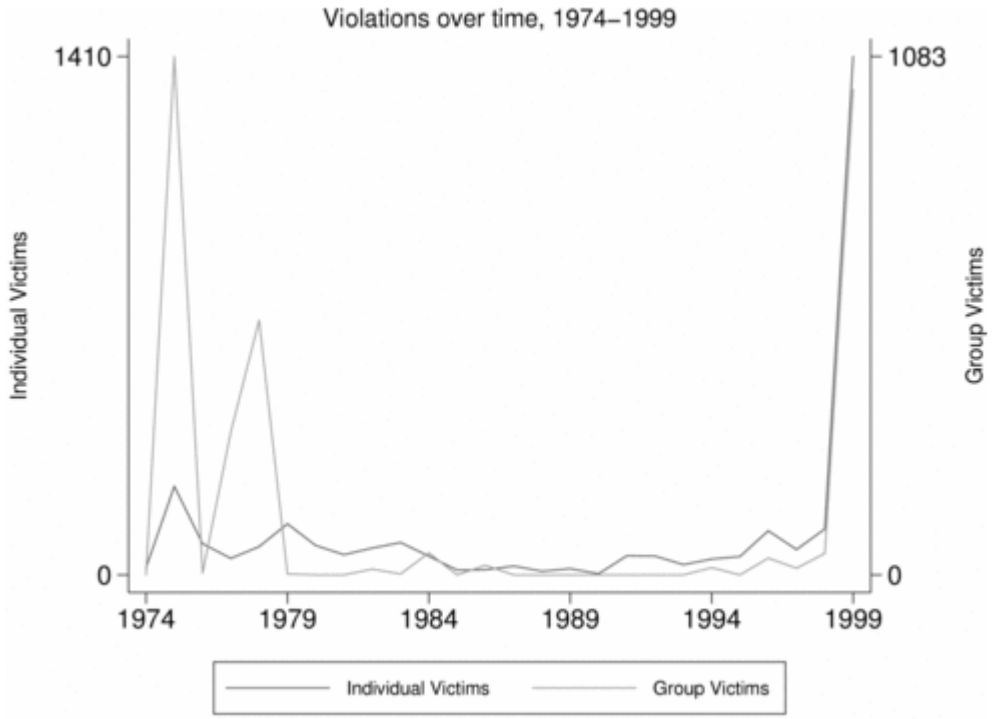
Figure 49

Source: Database of Narrative Statements Given to the CAVR

Figure 50

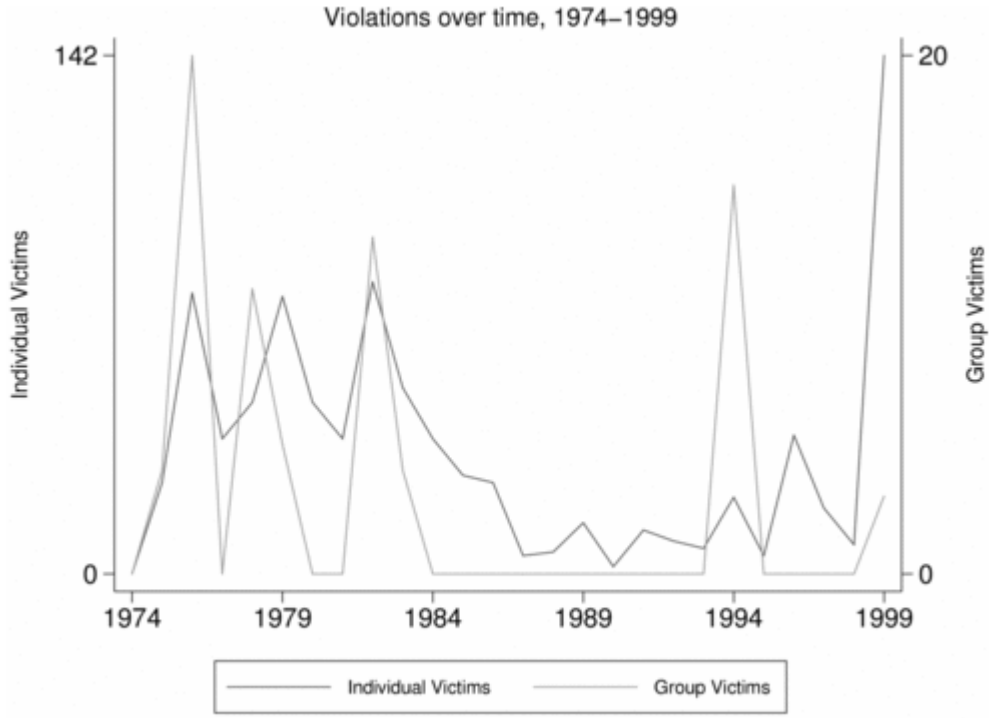
Source: Database of Narrative Statements Given to the CAVR

Figure 51



Source: Database of Narrative Statements Given to the CAVR

Figure 52



Source: Database of Narrative Statements Given to the CAVR

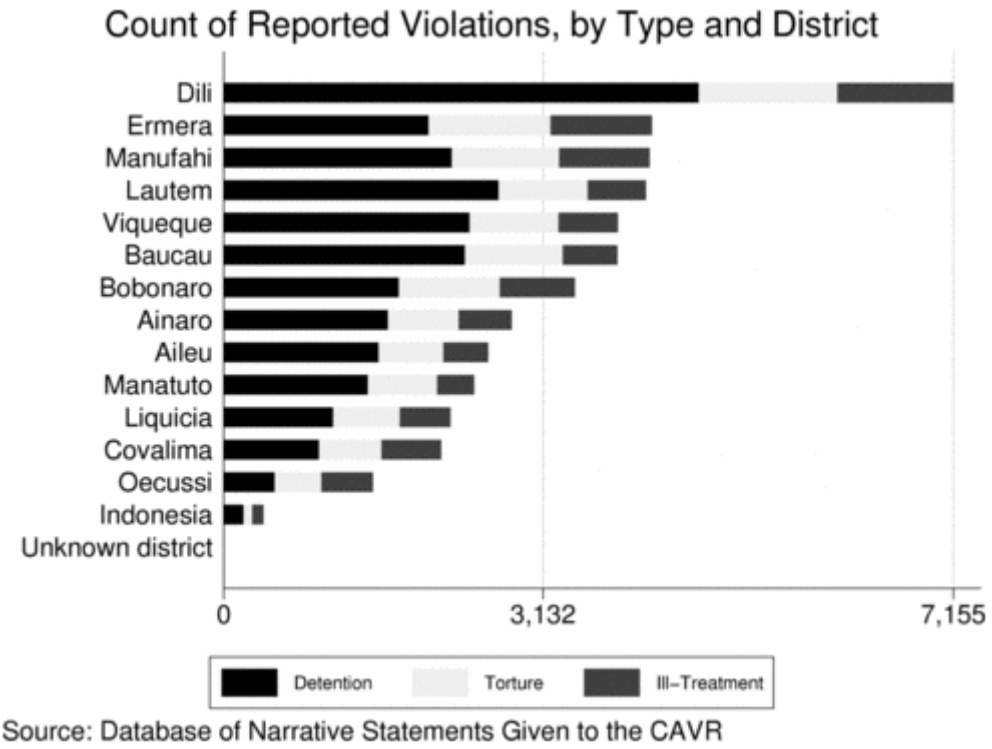
2.3.9 The Use of Detention and the Nature of Violations Committed During Detention Periods

Throughout the Indonesian occupation of Timor-Leste arbitrary detentions and displacement were employed throughout the territory. As the phenomenon of displacement is discussed in Section 1.3.7, this section focuses on the nature and use of detention during the Commission's mandate period.

2.3.9.1 The Relationship between Detention and Forms of Physical Abuse

In general reported acts of detention, torture and ill-treatment appear to be positively correlated. Detentions often occurred in the same events with physical abuse throughout the territory. This is reflected in Figure 53. The total number of reported acts of detention, torture and ill-treatment in Dili were higher than in any other district because the major detentions centers on Atauro Island and in the Comarca, Balide were both located in Dili.

Figure 53



Source: Database of Narrative Statements Given to the CAVR

Furthermore, reported detentions and torture over time are strongly positively correlated.[60] Figure 54 also shows that over time violence became increasingly coordinated and the magnitude of reported acts of torture increased over time (between the late 1970s and mid-1980s) relative to the number of reported detentions. This pattern might reflect the perpetrators' increasing capacity to target specific victims as the Indonesian occupation moved from its preliminary phase in the late 1970s and early 1980s to the consolidation phase from 1985 onwards. In the early invasion years there are approximately three reported cases of detention for each reported case of torture. After 1985, the two violations appear to be more closely linked, with approximately the same number of reported detentions and reported acts of torture each year. The resulting statistical pattern suggests that over time (and particularly after 1984) the practice of arbitrary detention became more targeted and was used more regularly in combination with acts of torture.

[60] The correlation coefficient between reported tortures and detentions by year between 1974 and 1999 is 0.81.

Figure 54

Source: Database of Narrative Statements Given to the CAVR

2.3.9.2 Patterns of Violations Committed During Periods of Detention

Of all the documented violations reported to the Commission during its narrative statement-taking process, detentions were the most frequently reported - representing 42.3% (25,383/60,047) of documented non-fatal violations. However, the use of detention was often combined with other forms of abuse: of the main forms of physical abuse reported to the Commission, at least 28.3% (7,174/25,383) were committed while the victim was held in detention. This empirical finding indicates that during detention victims were often vulnerable to other forms of physical abuse. This section explores the patterns of non-fatal forms of physical abuse committed during periods of detention and those committed while the victim was not detained.

The Commission's information on detentions and non-fatal violations often contains imprecise location and/or date information. In particular, 33.9% (20,334/60,047) of non-fatal violations were missing information about the month and day on which the violation occurred, while 52.9% (31,739/60,047) were missing information about the day on which the violation occurred. 2.3% (1,379/60,047) of non-fatal violations were missing information about the sub-district in which the violation occurred, while 31.2% (18,722/60,047) of non-fatal violations were missing information about the suco in which the violation occurred. Hence, the following analysis of forms of physical abuse and their relationship to the victim's status as a detainee is limited by the lack of precise dates and locations in the reported data.

Some forms of physical abuse were reported to have occurred more frequently in detention than others. In particular, Table 19 shows that, the abuses which were most often committed during known periods of detention were torture (38.4%, 4,267/9,094), ill-treatment (33.2%, 27,998/9,094) and threats (21.3%, 634/9,094). Furthermore, torture and ill- treatment are

reported much less frequently among victims who never have been held in detention: of the torture violations documented by the commission, 16.4% (1,820/11,123) were suffered by victims who never experienced detention. Of the acts of ill-treatment documented by the Commission 26.4% (2,227/8,436) were suffered by victims who never experienced detention. This is suggestive of the increased vulnerability of victims who are held in detention to being subjected to torture or ill-treatment.

The statistical data alone cannot clarify whether the association between detention and physical abuse was part of a formal policy by perpetrators to combine physical abuse with arbitrary detention, or whether the correlation reflects opportunistic behavior by military, police and other officials. However, the Commission's qualitative and historical research is informative in this regard. The Commission's qualitative research has identified evidence of policy and practice which encouraged the use of detention and special interrogation methods.

Table 19: Reported Non-Fatal Violations & Their Detention Context by Violation Type, 1974-1999

Violation Type	Victim detained at time of violation	% detained at time of violation	Victim detained but unknown dates	% detentions with unknown dates	Victim detained - but not at time of violation	% detained - but not at time of violation	Victim Never detained	% never detained	Total
Torture	4,267	38.4	4,569	41.1	1,820	16.4	467	4.2	11,123
Ill-Treatment	2,798	33.2	3,061	36.3	2,227	26.4	350	4.1	8,436
Threats	634	21.3	723	24.2	1,442	48.4	183	6.1	2,982
Forced Recruitment	166	7.7	851	39.5	1,049	48.6	91	4.2	2,157
Sexually-Based Violations	109	12.8	354	41.5	367	43.0	23	2.7	853
Property/ Economic Violations	313	6.6	810	17.1	3,355	70.9	257	5.4	4,735
Other	807	18.6	1,661	38.3	1,631	37.6	240	5.5	4,339
Total	**9,094**	**100.0**	**12,029**	**100.0**	**11,891**	**100.0**	**1,611**	**100.0**	**34,625**

Source: Database of Narrative Statements Given to the CAVR

The cases documented by the Commission indicate a change in the relationship between non-fatal violations and detention as the conflict moved into its last phase. As Table 20 shows, 56.7% (5,592/9,855) of non-fatal violations in 1999 were committed against victims who were not in detention and had never been detained before. In 1999 the proportion of reported non-fatal violations which were committed outside places of detention is more than double that for the two

earlier phases. Hence, it appears that in the first two phases of the conflict the use of detention had a stronger association with the commitment of other non-fatal violations.[61]

Table 20: Reported Violations & Their Detention Context by phase, 1974 - 1999

Phase Of Conflict	Victim detained at time of violation	% detained at time of violation	Victim detained but unknown dates	% detentions with unknown dates	Victim detained - but not at time of violation	% detained - but not at time of violation	Victim Never detained	% never detained	Total
Phase 1 (1974-1983)	2,963	18.8	8,006	50.8	4,357	27.6	446	2.8	15,772
Phase 2 (1984-1998)	3,407	37.9	3,011	33.5	1,942	21.6	638	7.1	8,998
Phase 3 (1999)	2,724	27.6	1,012	10.3	5,592	56.7	527	5.3	9,855
Total	**9,094**	**100.0**	**12,029**	**100.0**	**11,891**	**100.0**	**1,611**	**100.0**	**34,625**

Source: Database of Narrative Statements Given to the CAVR

Of the reported violations committed during a known period of detention, 16.5% (505/9,094) occurred in Dili. Hence, relative to other districts, a notably higher proportion of the reported violations occurring in detention were committed in Dili. While 42.8% (695/1,623) of violations in Liquiça were reported to have been committed during periods of detention, and 41.5% (886/2,135) in Covalima, the figure for Dili is 40.3% (1,504/3,731). Whereas, of the non-fatal violations suffered by persons in Oecusse and Indonesia, 76.0% (2,209/2,897) and 73.2% (390/533) respectively were suffered by victims who have never been detained. Hence, districts which reported relatively higher proportions of torture and ill treatment, tended to also report higher proportions of abuse within detention.

[61] It is difficult to make conclusive findings about the relative magnitude of non-fatal violations committed in detention in Phases 1 and 2 of the conflict, given that 50.8% (8,006/15,772) of detentions during Phase 1 and 33.5% (3,011/8,998) of detentions during Phase 2 lack sufficiently precise date information to determine whether they are associated with other violations suffered by the victim.

Table 21: Reported Non-Fatal Violations & Their Detention Context by Geographic Location, 1974-1999

District	Victim detained at time of violation	% detained at time of violation	Victim detained but unknown dates	% detentions with unknown dates	Victim detained - but not at time of violation	% detained - but not at time of violation	Victim Never detained	% never detained	Total
Lautem	455	19.7	1,307	56.6	479	20.8	67	2.9	2,308
Viqueque	401	13.0	1,371	44.4	1,245	40.4	68	2.2	3,085
Baucau	785	31.8	891	36.1	639	25.9	155	6.3	2,470
Manatuto	271	16.0	921	54.5	465	27.5	34	2.0	1,691
Manufahi	771	23.8	1,305	40.3	924	28.5	237	7.3	3,237
Aileu	477	26.6	867	48.3	412	23.0	38	2.1	1,794
Ermera	986	24.8	1,128	28.4	1,641	41.3	216	5.4	3,971
Liquiçia	695	42.8	448	27.6	405	25.0	75	4.6	1,623
Dili	1,504	40.3	1,267	34.0	646	17.3	314	8.4	3,731
Ainaro	457	21.6	1,005	47.5	582	27.5	72	3.4	2,116
Covalima	886	41.5	401	18.8	729	34.1	119	5.6	2,135
Oecussi	366	12.6	249	8.6	2,201	76.0	81	2.8	2,897
Bobonaro	992	32.7	793	26.1	1,133	37.3	116	3.8	3,034
Indonesia	48	9.0	76	14.3	390	73.2	19	3.6	533
Total	**9,094**		**12,029**		**11,891**		**1,611**		**34,625**

Source: Database of Narrative Statements Given to the CAVR

Male victims experienced the bulk of detentions, constituting 85.6% (21,273/25,383) of the total. Relative to female victims, males were twice as likely to be subjected to another violation during their detention. As Table 22 shows, of the male victims of detention, at least 28.1% (8,323/29,599) suffered another non-fatal violation, compared with 14.8% (716/4,833) for females who suffered another violation while detained.

Table 22: Reported Non-Fatal Violations & Their Detention Context by Sex of Victim, 1974 - 1999

Victim's Sex	Victim detained at time of violation	% detained at time of violation	Victim detained but unknown dates	% detentions with unknown dates	Victim detained - but not at time of violation	% detained - but not at time of violation	Victim Never detained	% never detained	Total
Female	716	14.8	1,671	34.6	2,304	47.7	142	2.9	4,833
Male	8,323	28.1	10,303	34.8	9,505	32.1	1,468	5.0	29,599
Unknown Sex	55	28.5	55	28.5	82	42.5	1	0.5	193
Total	**9,094**		**12,029**		**11,891**		**1,611**		**34,625**

Source: Database of Narrative Statements Given to the CAVR

According to the non-fatal violations data documented by the Commission, people of different ages suffered different levels of abuse while in detention. In particular, of those victims who suffered a known non-fatal violation during a period of detention 55.5% (5,044/9,094) were young or middle-aged adults (that is, persons between 15 and 49). Children and older people were detained substantially less often, and when they were detained, they were subjected to proportionally lower levels of abuse.

Table 23: Reported Non-Fatal Violations & Their Detention Context by Victim's Age Group, 1974 - 1999

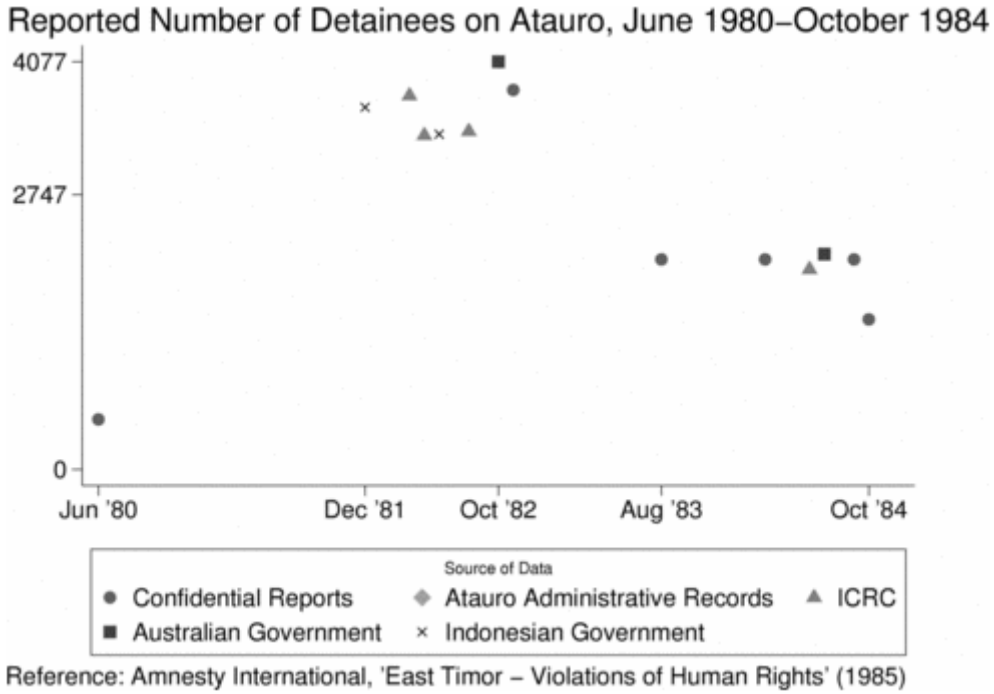
Victim's Age Group	Victim detained at time of violation	% detained at time of violation	Victim detained but unknown dates	% detentions with unknown dates	Victim detained - but not at time of violation	% detained - but not at time of violation	Victim Never detained	% never detained	Total
0-4	89	15.4	202	34.9	277	47.8	11	1.9	579
5-9	20	10.9	45	24.6	113	61.7	5	2.7	183
10-14	62	11.7	219	41.2	241	45.4	9	1.7	531
15-19	384	26.4	564	38.7	451	31.0	58	4.0	1,457
20-24	942	32.4	1,070	36.8	725	24.9	171	5.9	2,908
25-29	1,080	34.4	924	29.4	960	30.6	174	5.5	3,138
30-34	1,058	31.8	,1049	31.5	1,049	31.5	173	5.2	3,329
35-39	719	28.2	789	30.9	889	34.8	154	6.0	2,551
40-44	564	26.4	704	33.0	755	35.4	111	5.2	2,134
45-49	315	23.0	512	37.4	450	32.9	91	6.7	1,368
50-54	235	19.7	513	43.1	402	33.8	40	3.4	1,190
55-59	82	17.4	176	37.3	188	39.8	26	5.5	472
60-64	93	20.9	188	42.2	147	33.0	17	3.8	445
65-69	32	17.7	78	43.1	67	37.0	4	2.2	181
70-74	18	9.7	95	51.4	64	34.6	8	4.3	185
75-79	12	19.7	23	37.7	24	39.3	2	3.3	61
80+	12	18.5	23	35.4	29	44.6	1	1.5	65
Unknown Age	3,377	24.4	4,855	35.1	5,060	36.5	556	4.0	13,848
Total	9,094		12,029		11,891		1,611		34,625

Source: Database of Narrative Statements Given to the CAVR

2.3.9.3 Reported Statistical Patterns of Detention on Atauro

While arbitrary detention was used throughout the conflict in Timor-Leste, detentions on Atauro Island were reported to have been used mainly between 1980 and 1984. This is consistent with information collected by the Commission directly through its statement-taking process and in the secondary source reports collected from various sources by Amnesty International. However, there is a considerable difference in the level of documented detentions between these two data sources. The sources gathered by Amnesty International suggest that the detainee population on Atauro grew from about 500 in mid-1980 to around 3,500 in mid-1982 before declining to around 1,500 in October 1984, as shown in Figure 55.

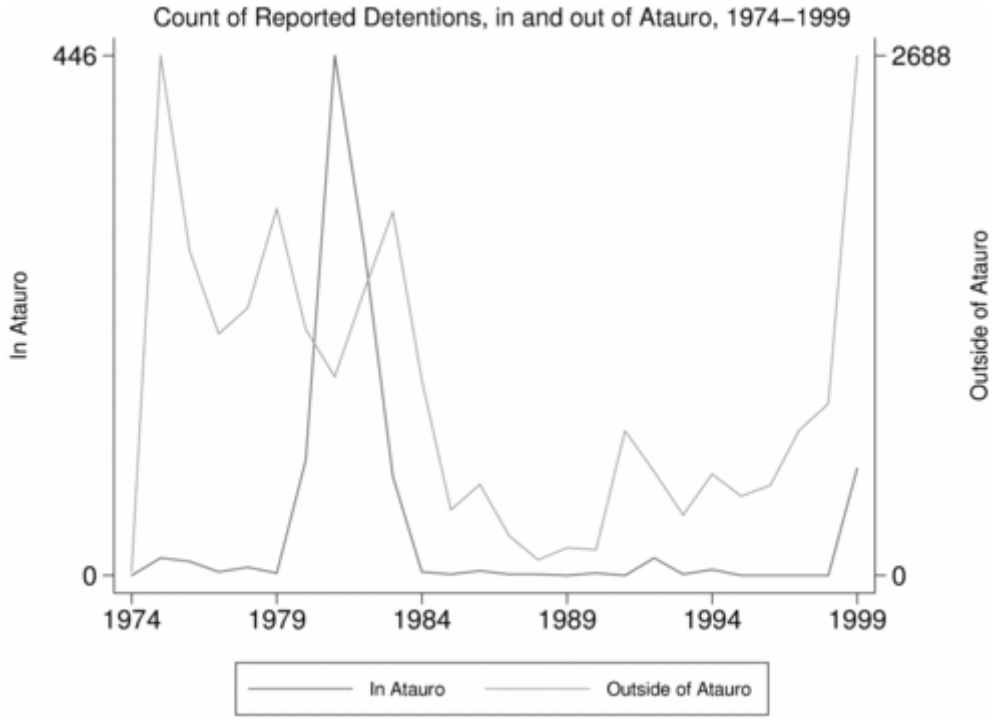
Figure 55



Reference: Amnesty International, 'East Timor – Violations of Human Rights' (1985)

According to the Commission's data, reported detentions on Atauro peak at 446 detainees in 1982, as is shown in Figure 56. Given that Amnesty's data was collected from multiple eyewitness accounts, including by the International Committee of the Red Cross, and Australian government officials, and from Indonesian administrative records, it is likely that the Commission's data significantly under-report detentions on Atauro. However, both the Amnesty and Commission data confirm that large groups of people were detained on Atauro Island in the early 1980s, in addition to continued large-scale detentions in other parts of Timor.

Figure 56



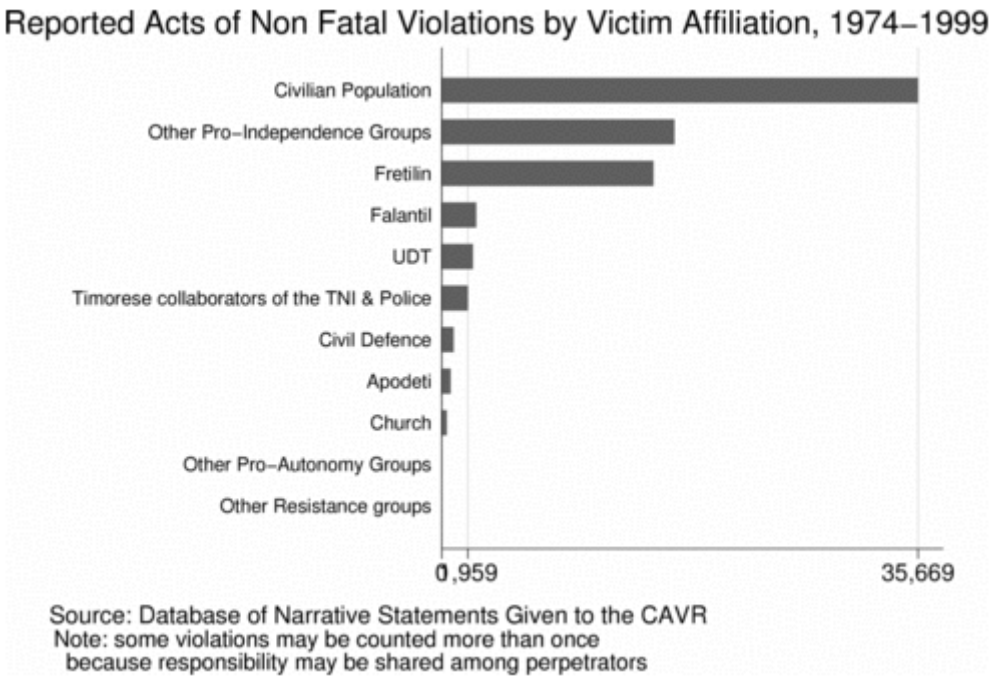
Source: Database of Narrative Statements Given to the CAVR

2.3.10 Patterns of violations by Political affiliations of Reported victims

A number of hypotheses examined by the Commission considered whether systematic and targeted campaigns were based on the victims' political affiliations. This section describes the reported extent and pattern of violations against civilians, armed-resistance fighters and political activists.

88.7% (68,943/77,748) of non-fatal violations reported to the Commission were violations against the civilian population, including both those civilians who were not known to have a political affiliation and those who were formally part of a pro-independence group or political party, as can be seen in Figure 57.

Figure 57



As the pro-independence movement grew more organized and popular in the lead-up to the UN-sponsored Popular Consultation in 1999, increasing numbers of civilians with pro-independence affiliations appear to have suffered non-fatal violations, as is seen in Figures 58, 59, 60, and 61.[62]

[62] It must be noted though, only 87 statements were collected in West Timorese refugee camps, and the Commission's district-based socialization process was convened in collaboration with local officials. As a result, people with pro-autonomy political affiliations may be under-represented in the Commission's statement-taking process.

Figure 58

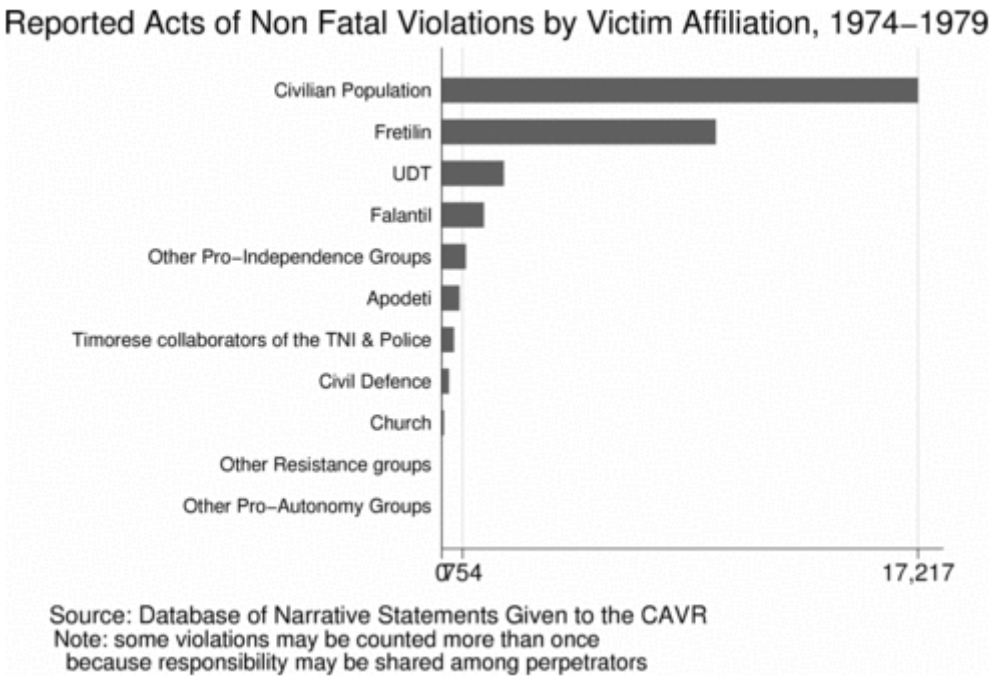


Figure 59

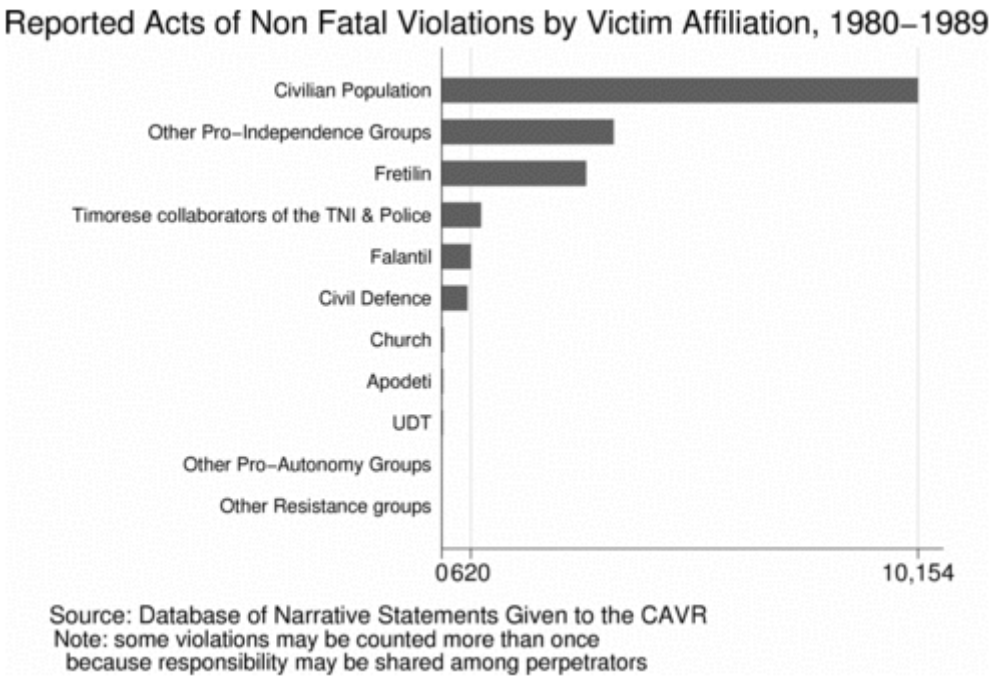


Figure 60

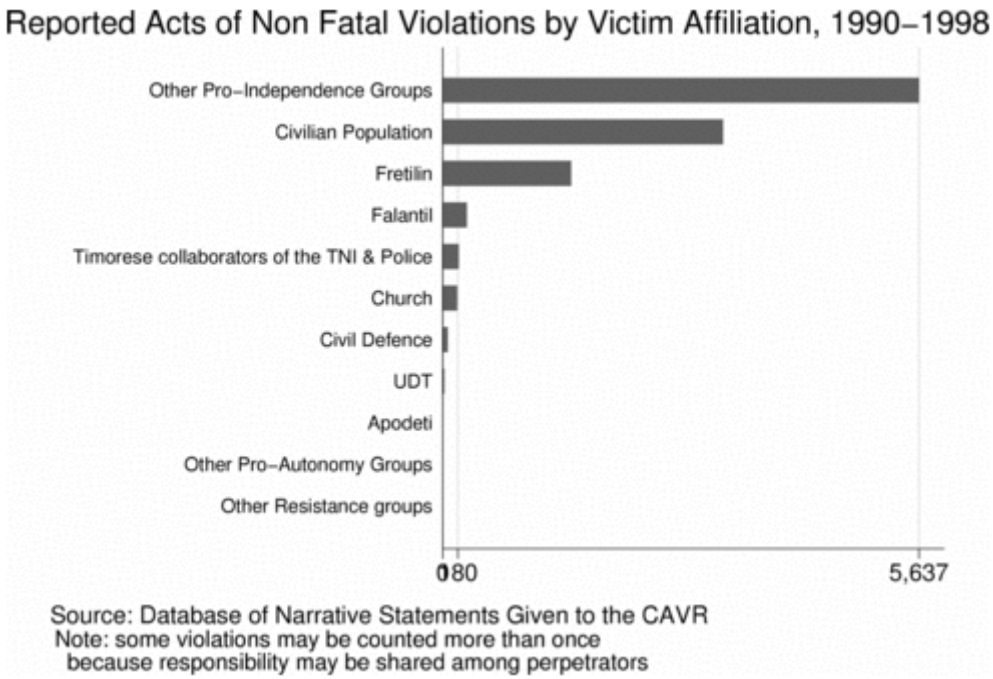
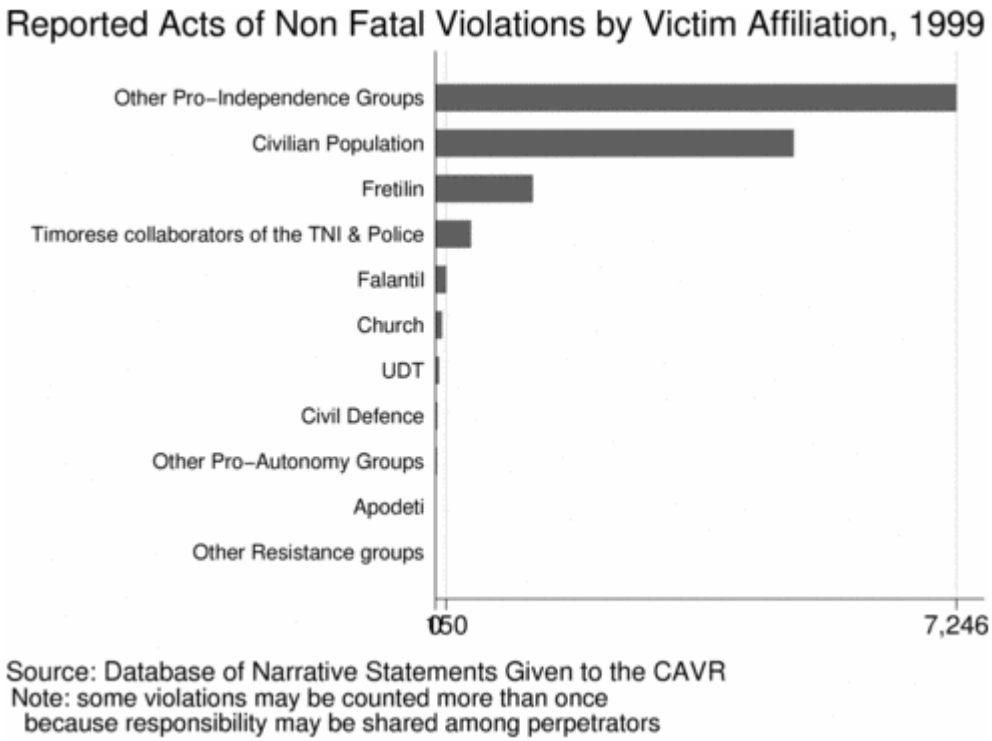


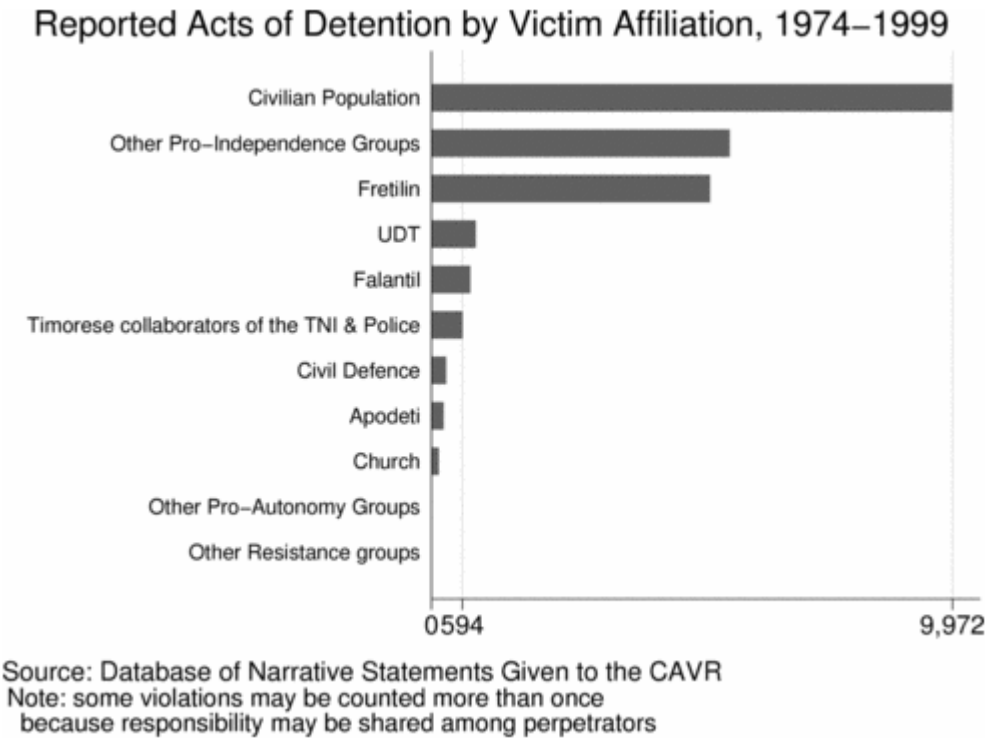
Figure 61



For the major non-fatal violation types (detention, torture, ill-treatment, forced labor, threats and property/economic violations), there are no substantial differences in the proportion share of documented victims by their political/social affiliations, as shown in Figures 62, 63, 64 and 65:

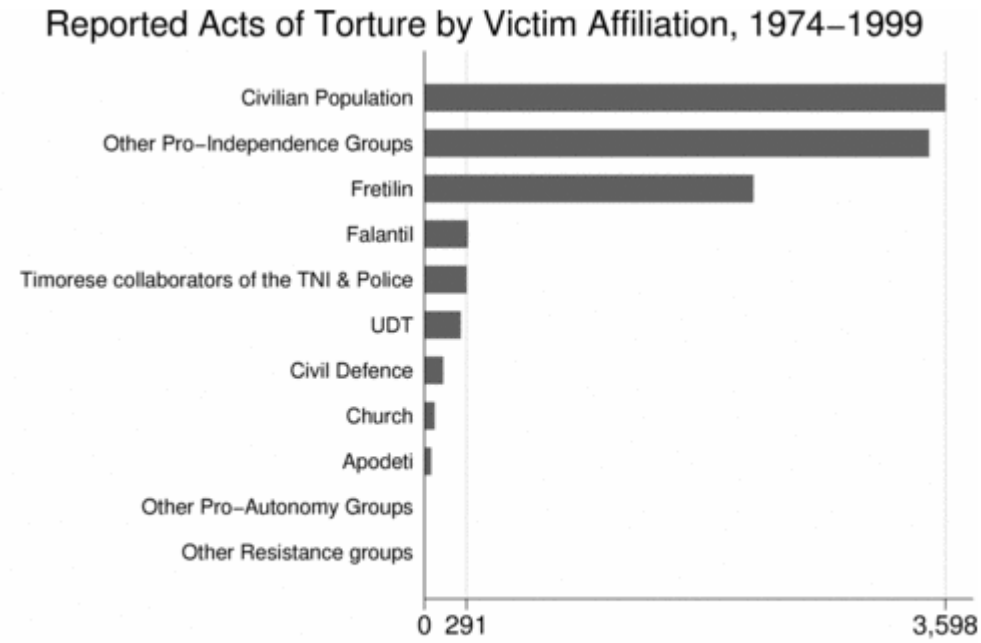
civilians without any political affiliations accounted for between 40-48% of the documented victims, whereas persons with a pro-independence affiliation accounted for between 43-55% of documented violations. For sexual violations, civilians without a known political affiliation accounted for a slightly higher proportion (56%, 441/770, of sexually based violations) than civilians known to be aligned with pro-independence groups and parties (43.1%, 427/770).

Figure 62



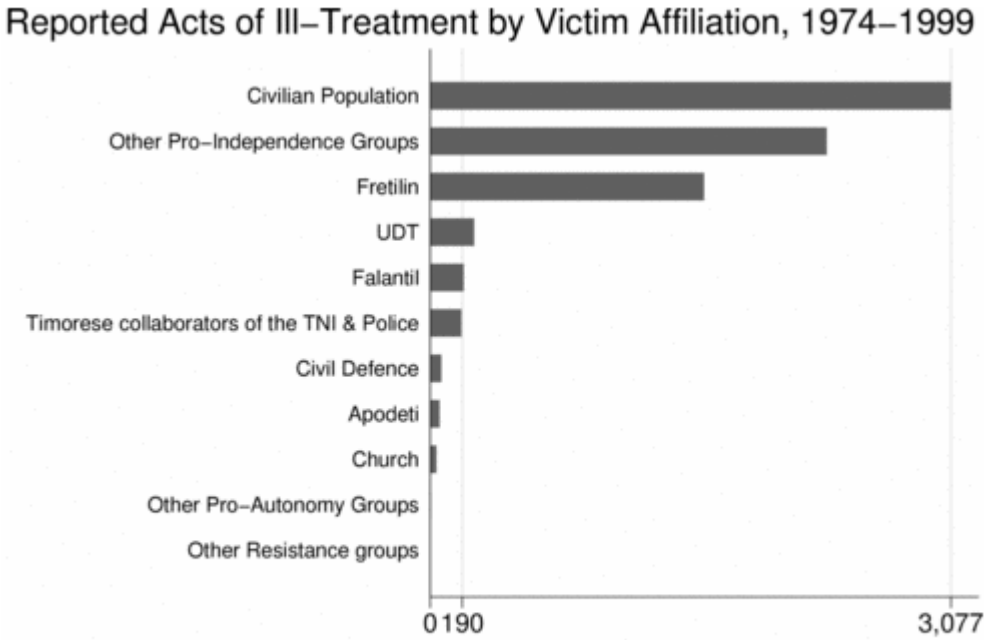
Source: Database of Narrative Statements Given to the CAVR
Note: some violations may be counted more than once because responsibility may be shared among perpetrators

Figure 63



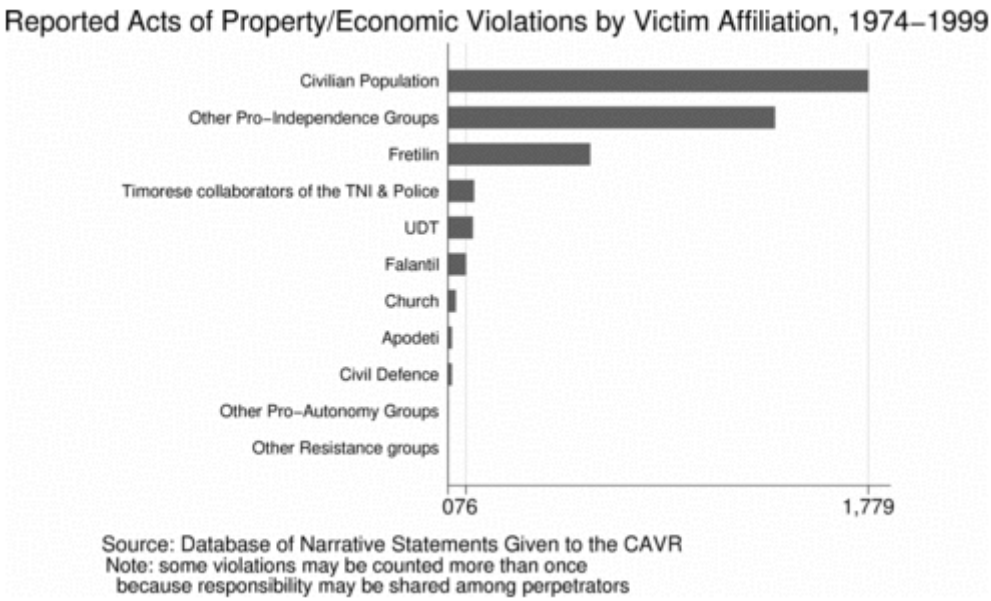
Source: Database of Narrative Statements Given to the CAVR
Note: some violations may be counted more than once
because responsibility may be shared among perpetrators

Figure 64



Source: Database of Narrative Statements Given to the CAVR
Note: some violations may be counted more than once
because responsibility may be shared among perpetrators

Figure 65



2.3.11 Reported Levels of Institutional Responsibility for Non-Fatal Violations

2.3.11.1 Overall Distribution of Attributed Institutional Responsibility for Reported Non-Fatal Violations

A number of different institutions were involved in acts of violence over the course of the conflict. The main institutional groups were the Indonesian military, FALINTIL, Timorese political parties (such as Fretilin, UDT and Apodeti), local administrative functionaries (such as Hansips and the Ciil Defense Forces) and militias. This section reviews the reported levels of responsibility for the main non fatal violations across the main institutional perpetrator groups.

The majority of non-fatal violations reported to the Commission were attributed to the Indonesian military and their Timorese collaborators, as shown in Table 24: 41.2% (37,298/90,635) of the perpetrator involvement in non-fatal violations was attributed to the Indonesian military, and 25.6% (23,230/90,635) to Timorese auxiliaries (such as the militias, civil defense force and local officials who worked under the Indonesian administration). For 7.9% (71,46/90,635) of reported violations, institutional perpetrator responsibility was attributed to the resistance groups and pro-independence forces.

Table 24: Count of Reported Non-Fatal Violation Type and Attributed Institutional Perpetrator, 1974-1999

Violation Type	Indonesian Military	Timorese Collaborators of TNI	Resistance Groups	Other	Civilian Population	Pro-Autonomy Groups	Un-known	Total
Detention	17,749	8,675	3,303	3,792	771	222	3,268	25,347
Torture	7,130	3,903	1,172	1,228	293	61	2,463	11,123
Ill-Treatment	4,628	3,354	1,075	1,252	214	96	2,216	8,436
Property/ Economic Violations	1,802	3,058	416	346	129	65	2,319	4,735
Other	2,367	1,634	789	735	73	34	1,020	4,339
Threats	1,458	1,590	236	251	55	28	1,143	2,982
Forced Recruitment	1,556	740	122	131	47	12	456	2,157
Sexually-Based Violations	608	276	33	26	11	1	203	853
Total	**37,298**	**23,230**	**7,146**	**7,761**	**1,593**	**519**	**13,088**	**59,972**

NOTE: Responsibility for violations may be shared among perpetrators, and therefore, columns may not be directly summed

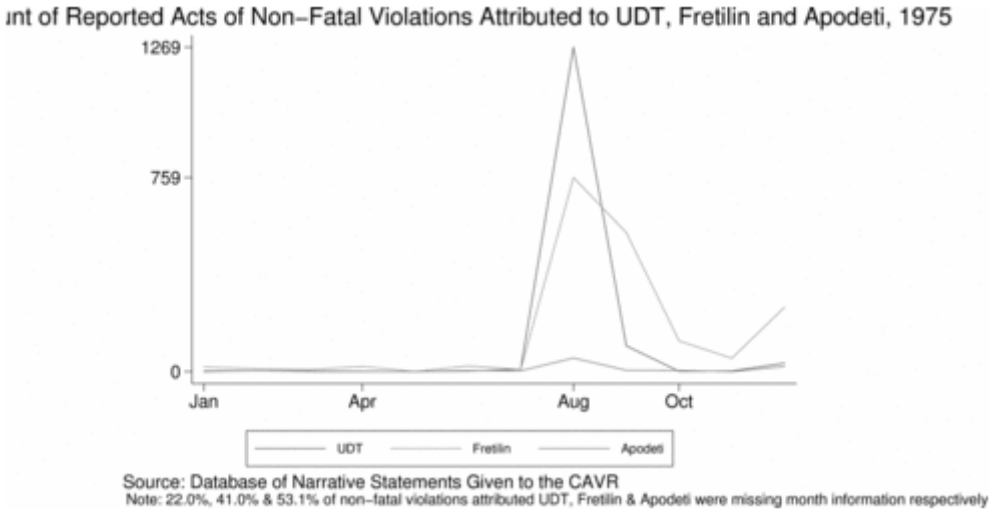
Source: Database of Narrative Statements Given to the CAVR

2.3.11.2 Temporal Patterns of Attributed Institutional Responsibility for Non-Fatal Violations

The levels of attributed institutional responsibility for documented non-fatal violations varied over the course of the conflict. During the initial Indonesian invasion in 1975, 51.0% (6,229/12,206) of perpetrator involvement in nonfatal violations documented by the Commission was attributed to the Indonesian Military, whereas 29.9% (3,653/12,206) to the Timorese resistance groups.

Of the documented non-fatal violations which occurred in 1975, 31.2% (3,169/10,162) were attributed to Fretilin, 19.4% (1,972/10,162) to UDT and 2.6% (261/10,162) to Apodeti. As is shown in Figure 66, the overwhelming majority of documented non-fatal violations in 1975 (where the exact month of the violation is known) attributed to Timorese political parties occurred in August and September.

Figure 66



As is shown in Figure 67, during the period in which the Indonesian military occupation developed from 1977 to 1984, the pattern of non-fatal violations attributed to the Indonesian Military and its Timorese auxiliaries is positively correlated.[63] Furthermore, a substantially higher relative proportion of perpetrator responsibility is attributed to the Civil Defense, Hansips and associated Timorese auxiliaries of TNI between 1977 and 1984 than during the initial invasion years (1975-1976) and consolidation years between 1985 and 1998. The Commission's statistical data is consistent with the hypothesis that the Indonesian military drew heavily on its Timorese auxiliaries between 1977 and 1984 in containing resistance activities and normalizing the occupation through physical integrity violations. Of the non-fatal violations attributed to Timorese auxiliaries between 1977 and 1984, 54.0% (4,660/8,633) were detentions, 16.6% (1,435/8,663) were tortures and 10.9% (938/8,633) were ill-treatments.

[63] The correlation coefficient for reported non-fatal violations attributed to the Indonesian military and police and those attributed to its Timorese auxiliaries is 0.88.

Figure 67

Source: Database of Narrative Statements Given to the CAVR

Deponents in the Commission's statement-taking process attributed substantial responsibility to specific units of the occupying Indonesian forces and their Timorese collaborators at particular times during the conflict. As Figures 68 and 69, show, after the party conflict and initial invasion in 1975, there is a relative peak in attributed responsibility of detentions and tortures by the Civil Defense between 1978 and 1983, after which the Kopassus (Special Forces Branch of the Indonesian military) carried out several hundred reported detentions and tortures in 1984 and 1986. In the late 1990's, as can be seen in Figure 70, responsibility is attributed to the police for detentions and tortures in the lead-up prior to the Popular Consultation. The reported magnitude of detentions and tortures attributed to the Civil Defense in 1983 is 1.6 times bigger than detentions and tortures attributed to Kopassus in the same year and 2.0 times bigger than those attributed to the police in 1999.

Figure 68

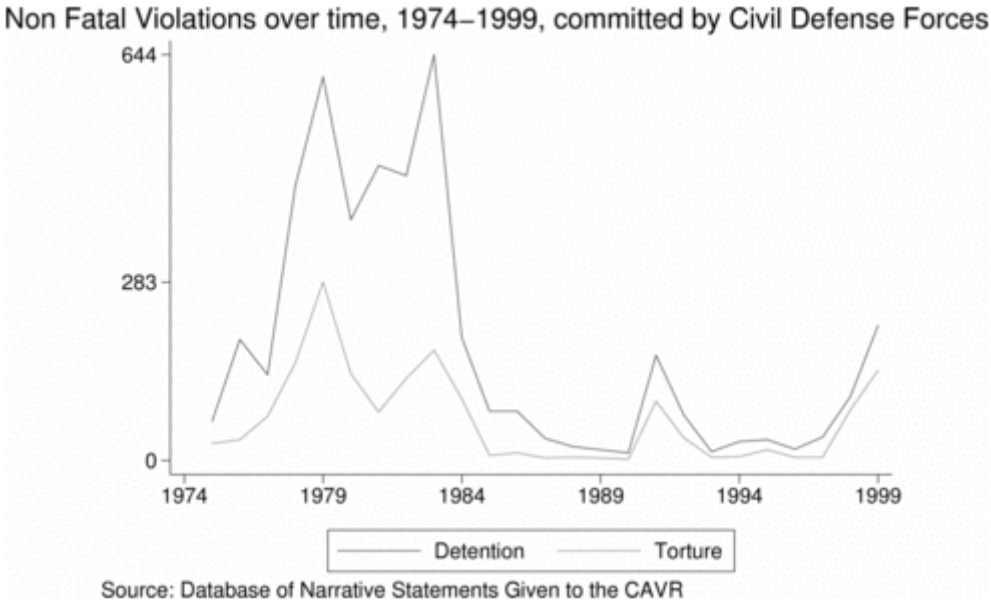


Figure 69

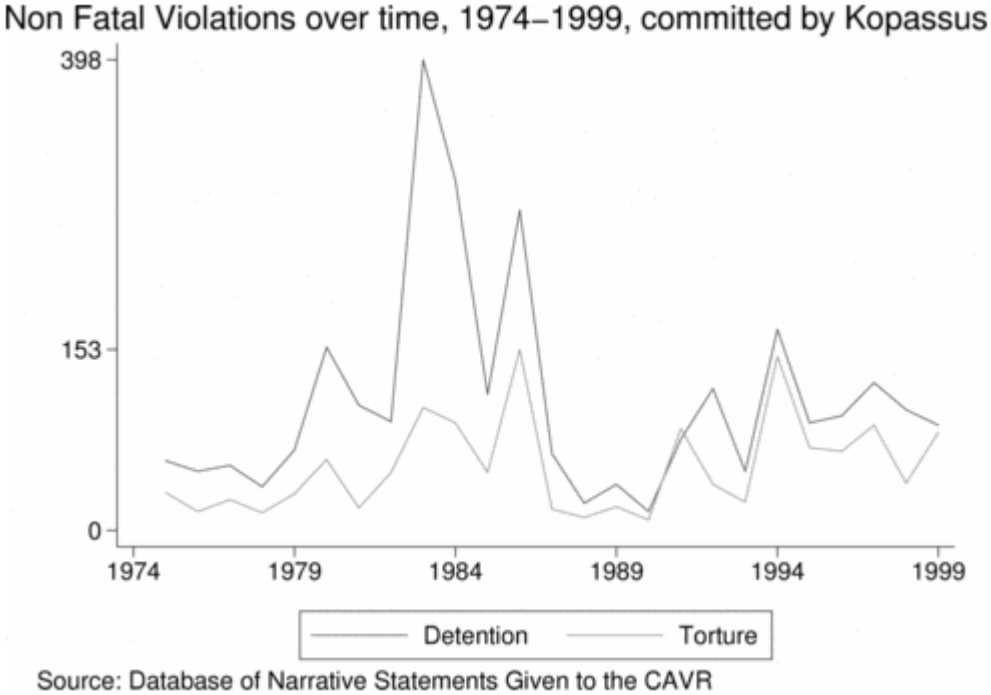


Figure 70

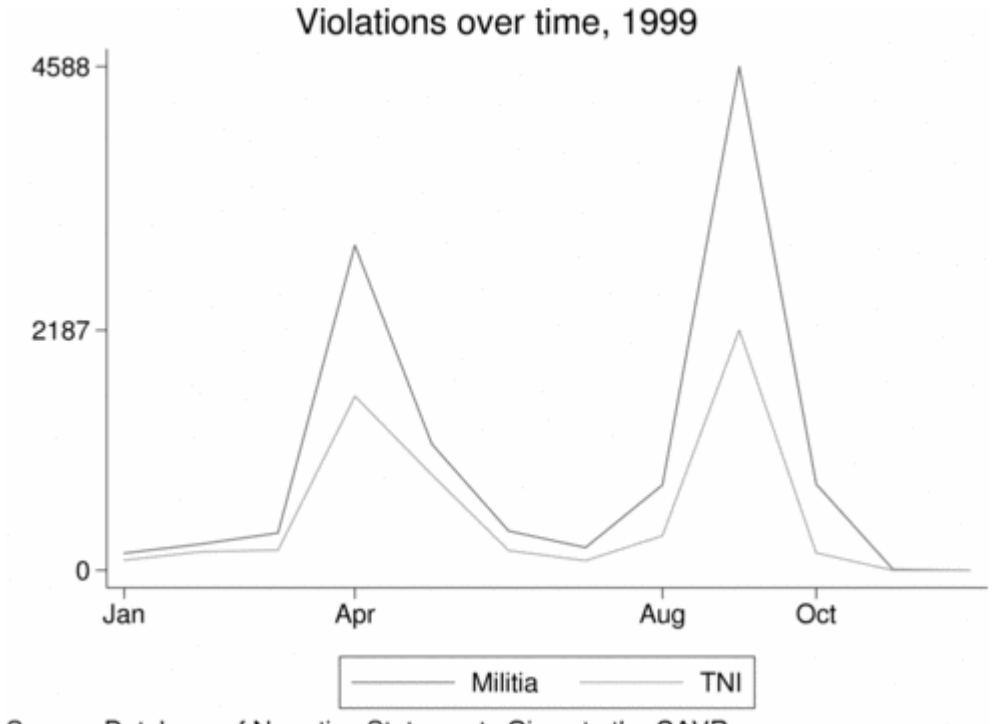
Source: Database of Narrative Statements Given to the CAVR

Whereas in 1999, overwhelming responsibility for non-fatal violations is attributed to the militias and Indonesian military – with the militias being associated with more than twice as non-fatal violations than the Indonesian military, as is shown in Figures 71 and 72.

Figure 71

Source: Database of Narrative Statements Given to the CAVR

Figure 72



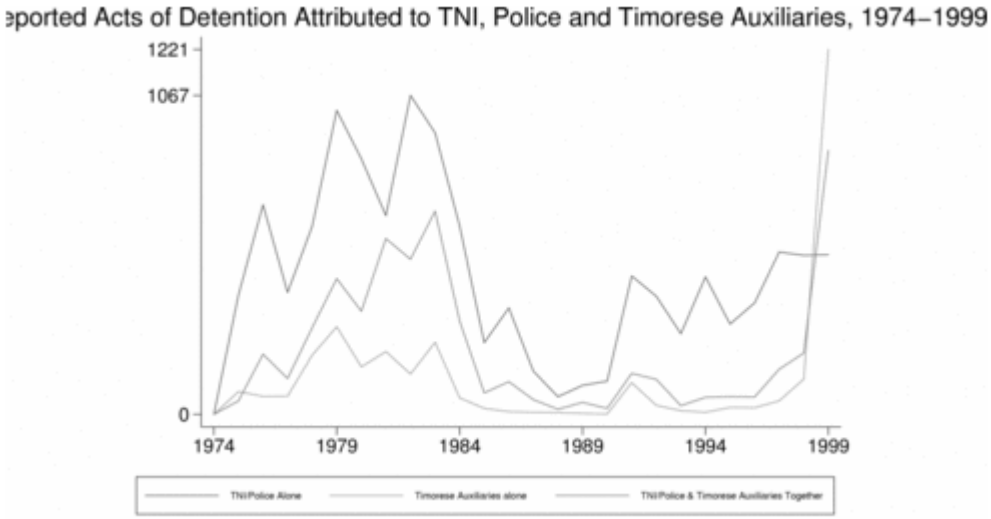
Source: Database of Narrative Statements Given to the CAVR

Sometimes non-fatal violations were attributed to the Indonesian military and police acting alone, other times to Timorese auxiliaries acting alone and other times to the Indonesian military

and police acting in concert with their Timorese auxiliaries. The pattern of shared and individual responsibility between the Indonesian forces and their Timorese auxiliaries differed by violation type and varied over time.

Of the acts of arbitrary detention documented by the Commission, 82.3% (20867/25347) were attributed to either the Indonesian occupying force and/or their Timorese auxiliaries. As Figure 73 shows, reported acts of arbitrary detention attributed to the Indonesian military and police alone, Timorese auxiliaries or both forces acting together are positively correlated over time. In particular, periods in which substantial documented acts of detention are attributed to both forces acting together (as well as each acting individually) include the period of consolidation and normalization of the occupation (particularly between 1978 and 1983) and around the time of the UN-sponsored Popular Consultation. Hence the Commission's quantitative analysis of arbitrary detentions is consistent with the hypothesis that coordination and cooperation between the Indonesian occupation force and their Timorese auxiliaries was particularly strong after the Indonesian military had secured large parts of Timor-Leste and started consolidating its occupation of the territory and then again in 1999 in the lead-up to and aftermath of the UN-sponsored Popular Consultation.

Figure 73

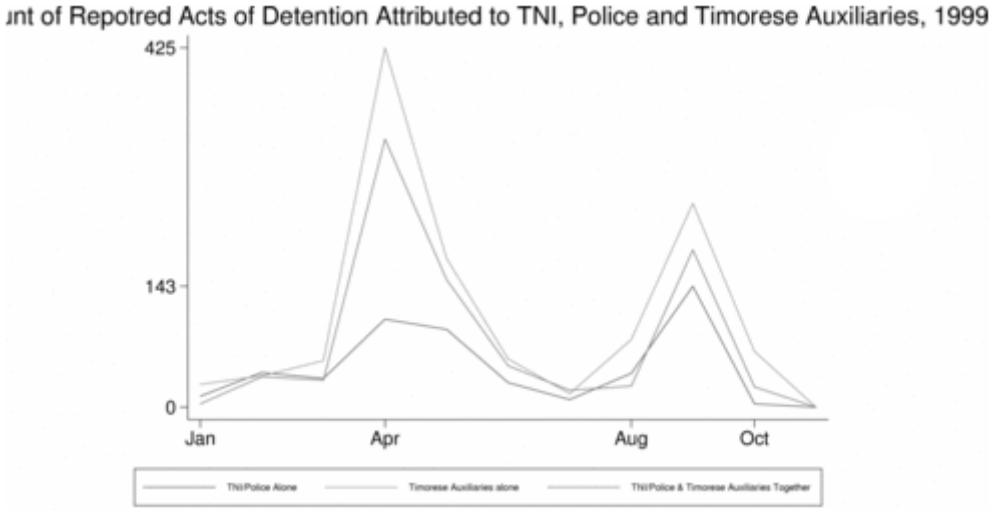


Source: Database of Narrative Statements Given to the CAVR

Between 1975 and 1998, substantially more acts of arbitrary detention are attributed to the Indonesian military acting alone relative to acts of detention which were solely attributed to Timorese auxiliaries or jointly to both the Indonesian occupying force and their Timorese auxiliaries. However, in 1999 most acts of detentions were attributed to Timorese auxiliaries. In particular, of the acts of arbitrary detention in 1999 documented by the Commission, 75.7% (2,104/2,779) were attributed to either the Timorese auxiliaries acting alone or in collaboration with the Indonesian military and police. Whereas, 19.2% (534/2,779) of documented acts of detention which occurred in 1999 were attributed to the Indonesian military alone. Almost all these acts were reported to have occurred in the months of April, May and September of 1999, as shown in Figure 74. The resulting statistical pattern is suggestive of prior planning and operational coordination between both forces in their use of arbitrary detention. Furthermore, during these months the Indonesian government was reassuring the United Nations that its

military was trying to bring the violence in Timor-Leste under control. Hence the Commission's statistical data is consistent with the hypothesis that in 1999 the Indonesian military and police aided and abetted their Timorese auxiliaries (principally the pro-autonomy militias) in the widespread use of arbitrary detention in the lead up to and aftermath of the UN-sponsored Popular Consultation.

Figure 74

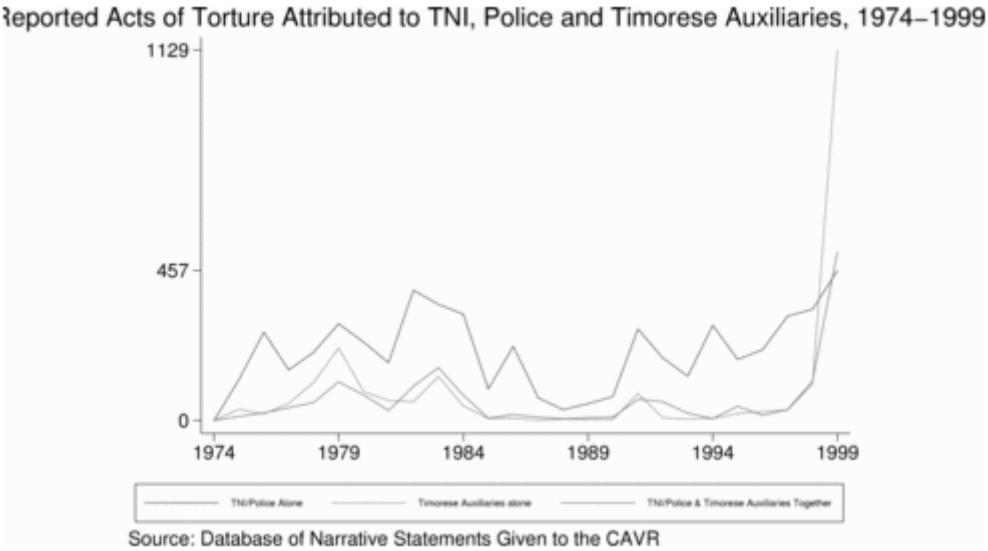


Source: Database of Narrative Statements Given to the CAVR

The pattern of responsibility attributed both solely and jointly to the Indonesian occupying forces and their Timorese auxiliaries has some notable similarities to arbitrary detentions, despite acts of ill-treatment and torture being used in a more targeted fashion.[64] A similar proportion (namely 82.5% (16,135/19,559) of the documented ill-treatments and tortures are attributed to the Indonesian occupying force and their Timorese auxiliaries. Also, similar to acts of arbitrary detention, in 1999 75.8% (3,278/4,324) of reported ill-treatments and tortures were attributed to be the responsibility of Timorese auxiliaries (either acting alone or in collaboration with associates of the Indonesian military and police). However a greater proportion of ill-treatments and tortures were reported to have occurred in 1999 than between 1974 and 1998 relative to documented cases of arbitrary detention, as can be seen when comparing Figures 75 and 76 to Figure 73.

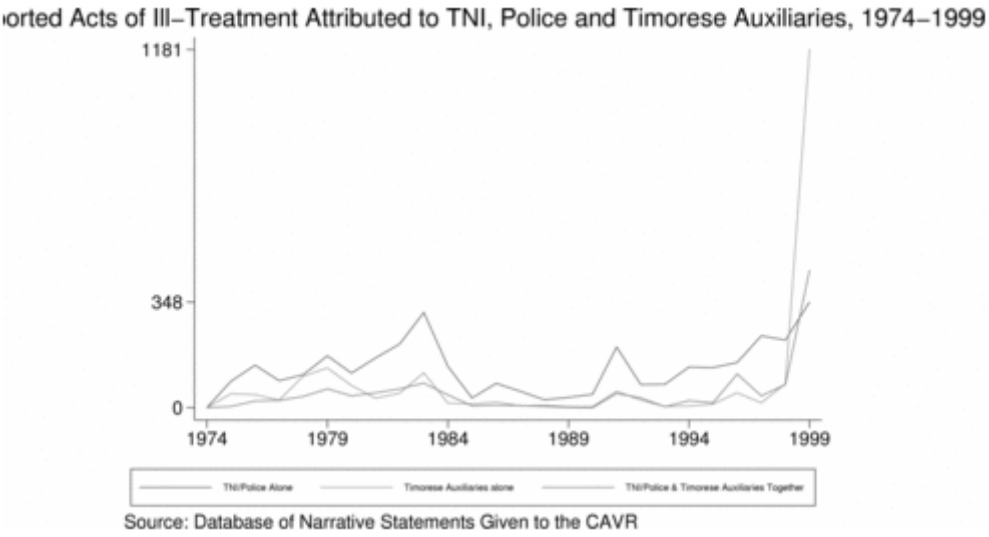
[64] For a detailed explanation on the more targeted nature of torture and ill-treatment relative to acts of detention, refer to Section 2.3.8 above.

Figure 75



Source: Database of Narrative Statements Given to the CAVR

Figure 76



Source: Database of Narrative Statements Given to the CAVR

The pattern and magnitude of documented tortures and ill-treatments in 1999, attributed to the Indonesian occupying forces and their Timorese auxiliaries both solely and jointly is similar to that of documented acts of detention in 1999, as can be seen in Figures 77 and 78.

Figure 77

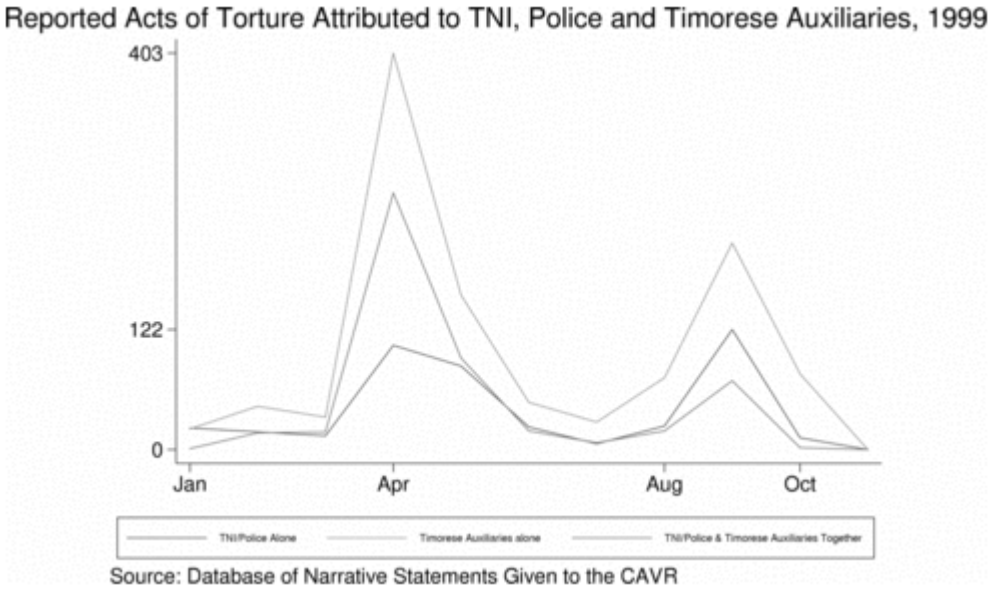
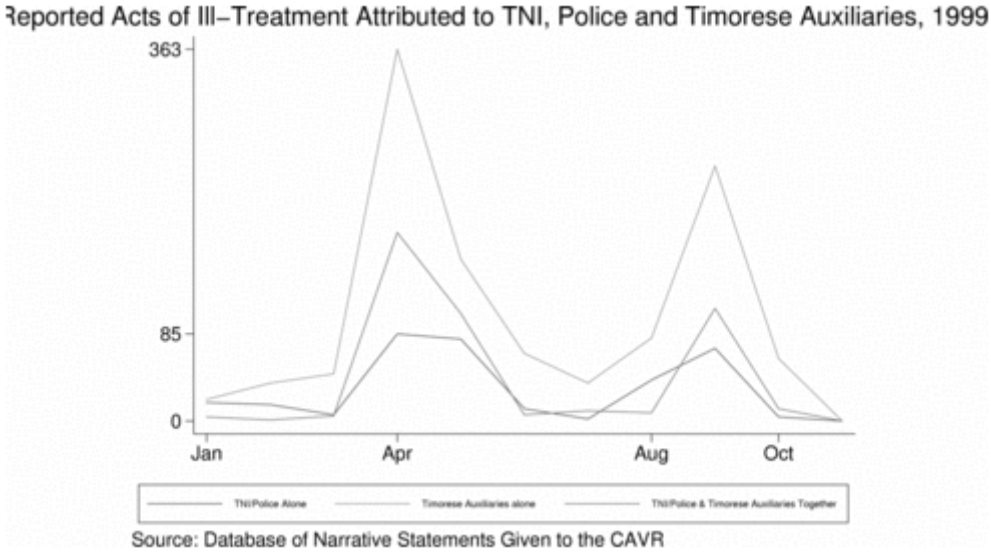


Figure 78

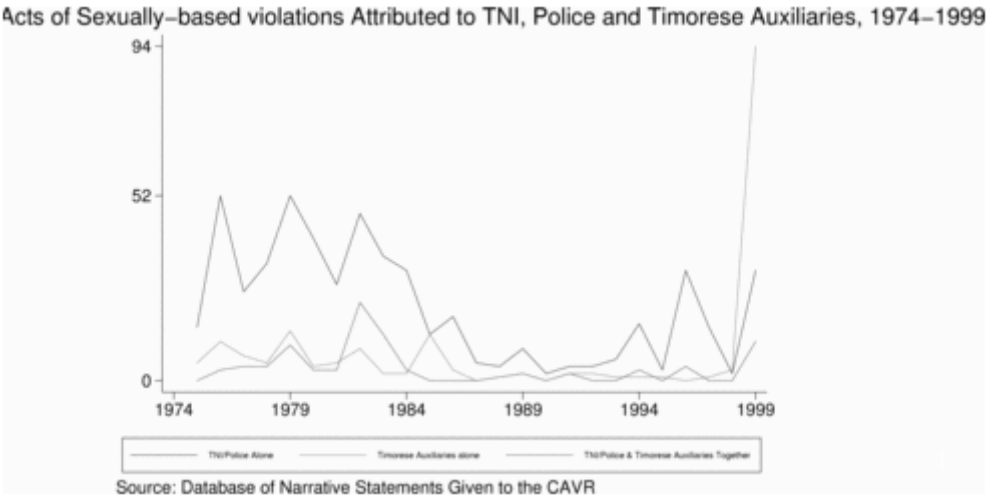


Hence as is the case for arbitrary detentions, the Commission's statistical data on ill-treatment and torture is consistent with the hypothesis that coordination and cooperation between the Indonesian occupation force and their Timorese auxiliaries was particularly strong after the Indonesian military had secured large parts of Timor-Leste and started consolidating its occupation of the territory and then again in 1999 in the lead-up to and aftermath of the UN-sponsored Popular Consultation.

The nature and pattern of attribution of perpetrator responsibility for documented sexually-based violations and property/economic violations were both notably different than for documented detentions, tortures and ill-treatments.

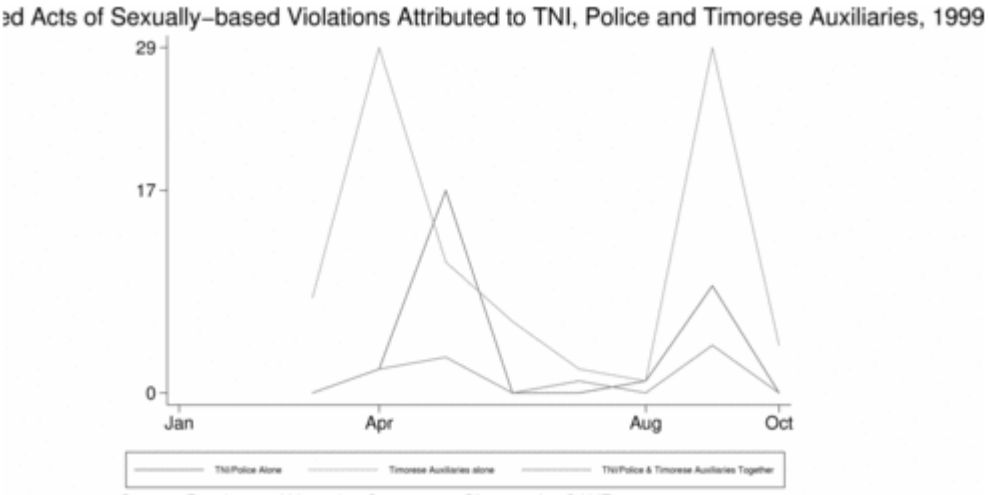
A higher proportion of sexually-based violations were attributed to the Indonesian military acting alone, whereas a much smaller proportion of sexually-based violations was attributed to the Indonesian occupying forces acting together with their Timorese auxiliaries. In particular, 61.0% (520/853) of documented sexually-based violations were attributed to the Indonesian military and police acting alone, 22.0% (188/853) to Timorese auxiliaries acting alone and 10.3% (88/853) to both forces acting together. As for all other non-fatal violations, the higher counts of sexually-based violations were attributed to Indonesian military alone between 1975 and 1998 than those attributed solely to Timorese auxiliaries or jointly to both forces. Whereas for 1999, the majority of sexually-based violations (66.2% (94/142)) reported to the Commission were solely attributed to the Timorese auxiliaries of the Indonesian military. These temporal patterns are shown in Figure 79 and 80.

Figure 79



Source: Database of Narrative Statements Given to the CAVR

Figure 80

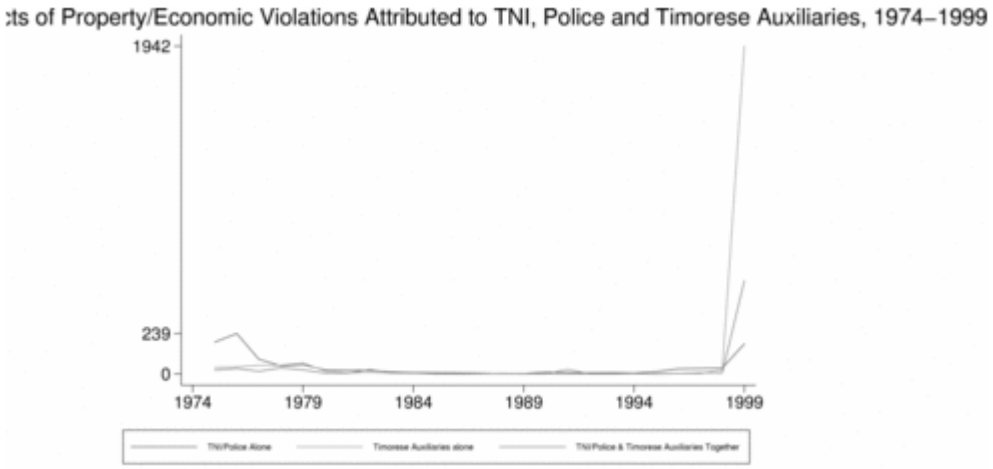


Source: Database of Narrative Statements Given to the CAVR

Of the documented property/economic violations attributed to the Indonesian occupying forces and/or their Timorese auxiliaries, 65.1% (2,673/4,105) occurred in 1999. As can be seen in

Figure 81, 70.2% (1,942/2,766) of the documented property/economic violations in 1999 were attributed solely to the Timorese auxiliaries of the Indonesian military and police, 20.0% (553/2,766) were attributed to both forces acting together and 6.4% (1,78/2,766) to the Indonesian military and police acting alone. The Commission's quantitative analysis is consistent with the hypothesis that most property and economic destruction was carried out in 1999 and was usually carried out by the militias acting alone or in collaboration with the Indonesian military and police.

Figure 81



Source: Database of Narrative Statements Given to the CAVR

3. Mauxiga Case-study: A Quantitative Analysis of Violations Experienced During Counter-Resistance Operations

3.1 Introduction

This section presents a detailed case study on the nature and pattern of violations experienced by the people of Mauchiga (Hatu Builico, Ainaro) in the early 1980s. The case-study takes the form of a descriptive statistical analysis based on data collected by two village leaders from Mauchiga.

3.1.1 Background to documentation effort

Over a period of 18 years village leaders from Mauchiga documented displacements, detentions and killings arising from a crackdown by the Indonesian military in connection with attacks organized by the Resistance in the area in August 1982. The Mauchiga Documentation Project was completed in August 2004 when village leaders handed over tabulated lists compiled from their narrative interviews to the Commission during a Public Hearing held by the Commission in Mauchiga.

The purpose of the project was to develop an accurate historical record of the extent, pattern, trend and nature of violations experienced by members of the community of Mauchiga during the early 1980s. Deponents were invited and encouraged to talk about any displacement, detention or fatal violation experienced by anyone they knew in connection with the August 1982 uprising (levantemanto).

3.1.2 Limitations of the data

The data on which this case-study are based were collected through a convenience sample of persons willing to report and share their experiences of human rights violations (namely displacements, arrests and detentions and conflict-related deaths) connected to the August 1982 uprising and counter-Resistance crackdown. Abilio dos Santos and Olga da Silva collected this data in two separate phases of data collection.[65] The first phase of data collection, from February 1986 until April 2003, involved Abilio dos Santos and Olga da Silva periodically carrying out narrative interviews in the different aldeias in village of Mauchiga. The two visited the following aldeias during their data collection and documentation work: Mauchiga, Hataquero, Goulora, Leotelo-1 and Leotelo-2. Deponents were selected based on the interviewers' own social networks and referrals by other interviewees. The second phase of data collection, from May 2003 to July 2004, consisted of the compilation of lists of victims of detention, displacement, and conflict-related deaths. During this second phase, a number of respondents who had given information in the first phase of data collection were re-interviewed to fill in gaps in the narrative information which had been collected.

The project restricted the information it collected to the specific violations of arbitrary detentions, displacement and conflict-related deaths which were directly connected to the events of 20 August 1982 and the crackdown that followed. It therefore did not document other forms of abuse, such as property destruction and sexual violence, nor did it document abuses which were connected to events other than those that occurred in connection with the 20 August uprising.

Provision was made for duplicate reporting on the same victim by multiple deponents at two stages in the documentation process. First, the data collection team periodically scanned their lists for duplicate reports of victims. Second, once the data was entered into an electronic database, computerized searches and analytical tests were conducted to identify names which could be duplicate reports of the same victim.[66]

3.1.3 Historical background

On 6 July 1982 members of FALINTIL and of the local clandestine movement began planning a series of attacks on Indonesian military posts in the area surrounding Mauchiga. An informer betrayed their plans to the Indonesian military. On 10 July members of the Indonesian army and Hansip from Hatu Builico began house-to-house searches in Goulora, Mauchiga and Hatuquero. They arrested over 30 people, including 13 who had attended the 6 July meeting. Those arrested were taken directly to the district military headquarters (Kodim) in Ainaro Town. Over the following days, the Indonesian military arrested more people, whom they brought to the Hatu Builico Sub-district headquarters (Koramil).

Despite the arrests, on 20 August 1982, at about 4:30am, FALINTIL together with a number of men from Dare and Mauchiga attacked several ABRI posts around Mauchiga, including the Dare Koramil. The same day ABRI soldiers and Hansip from the posts that had been attacked as well as other posts in the area retaliated. During the following days additional ABRI troops from

[65] Abilio dos Santos is the Village Secretary of Mauchiga. Olga da Silva is a teacher at the primary school in Mauchiga. The Commission has chosen to follow the official RDTL spelling of "Mauchiga", though it is known to many also as "Mauxiga."

[66] This process uncovered seven duplicated reports of fatal violations.

outside the area, including units of Battalions 745 and 746, were also deployed around Mauchiga. Between 20 August and 24 August Indonesian troops and Hansip destroyed and looted property. A large proportion of the population of the village of Mauchiga were either forcibly displaced or fled the village out of fear for their safety. The Indonesian military forcibly transferred villagers to several different locations, including the island of Ataúro (Dili), Dotik (Alas, Manufahi) and Dare (Hatu Builico, Ainaro).

3.2 Descriptive statistical analysis of violations reported to Mauchiga Documentation Project

3.2.1 Reported displacements and detentions suffered by Mauchiga residents

3.2.1.1 The demographic profile of victims of reported displacements and detentions

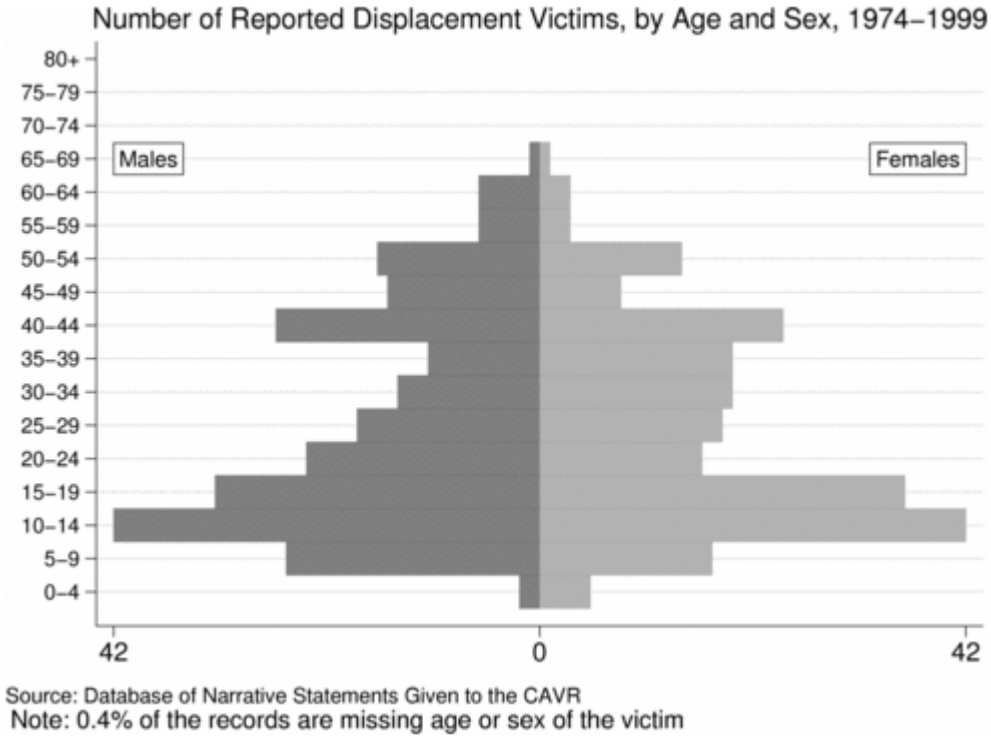
The project documented 1,803 acts of displacement involving Mauchiga residents between July 1982 and January 1986. These 1,803 displacements were experienced by 464 individual residents of Mauchiga: 48.7% (226/464) of whom were females and 38.8% (180/464) were children.[67]

These 464 victims of displacement and detention amounted to approximately 20.4% (464/2,269) of the total population of Mauchiga Village.[68] Hence, the project's findings are consistent with the hypothesis that displacement was widespread in Mauchiga during the 1980s.

Of the reported victims of displacement, 80.0% (371/464) of reported victims were initially arrested and displaced with their families. The remaining 20% (93/464) of documented victims of displacement were initially detained by themselves (and not along with their families).[69] As Figure 82 shows, 41.2% (191/464) of the displaced were between the ages of 10 and 24. As was and still is the case in most parts of Timor-Leste, the population of Mauchiga was over-represented by persons under 25. Hence the findings of the Mauchiga Documentation Project are consistent with the hypothesis that the Indonesian military actively sort to eliminate the social and operational base of the resistance movement in Mauchiga by forcibly deporting the general population (including women, children and the elderly).

[67] The Commission used the internationally-recognized standard that defines children as persons under the age of 18 (see Article 1of the UN Convention on the Rights of the Child (UN Doc A/44/49 (1989)), which was adopted by General Assembly Resolution 44/25 of 20 November 1989 and. entered into force on 2 September 1990.
[68] The 2001 Timor-Leste Suco Survey was used for the population base for Mauchiga.
[69] Those who were deported as individuals were displaced in this manner, as they were captured by the Indonesian military while they were alone, usually either in their agricultural gardens, or in other places away from their place of residence.

Figure 82



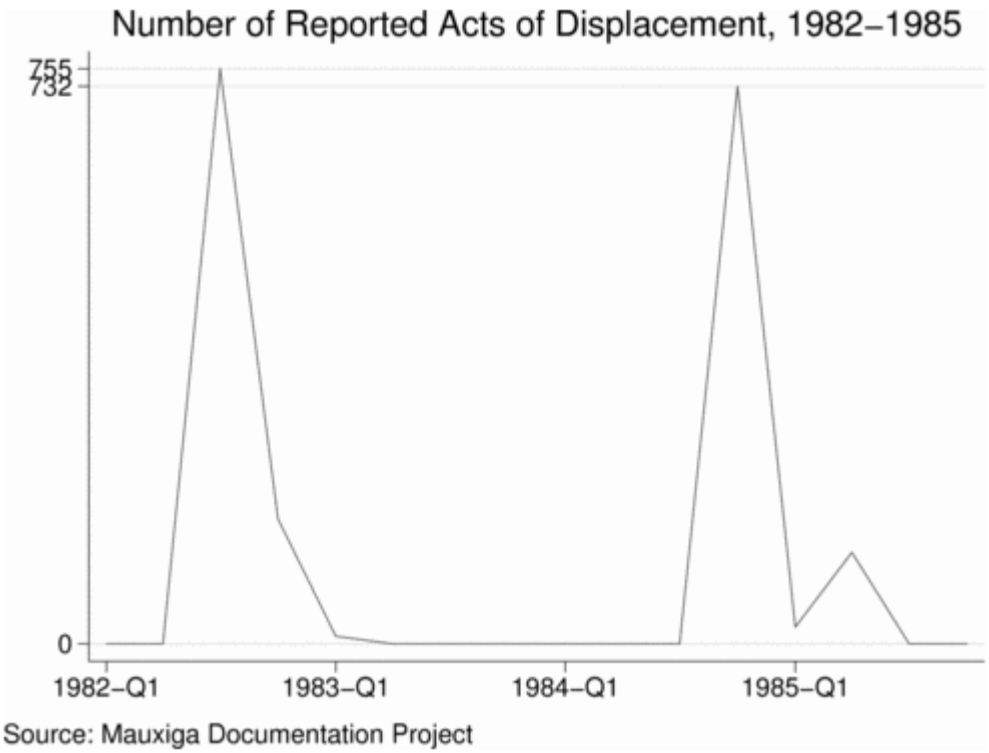
3.2.1.2 Accountability for the large-scale displacement and acts of detention of Mauchiga residents

All of the reported acts of displacement of these 1,803 individuals, documented by the Mauchiga Documentation Project, were attributed to the Indonesian military. In some specific cases, deponents reported the involvement of specific units of the Indonesian military or of the civil defense forces. The Hatu Builico Hansip were reported to have taken part in 31.7% (571/1,803) of reported acts of displacement involving Mauchiga residents. In these same acts of displacement it was reported to the Mauchiga Documentation Project that the Hatu Builico Hansip were working in collaboration with and under the direction of the Sub-Regional Command (Korem) in Dili, the District Military Command (Kodim) in Ainaro and the Hatu Builico Sub-district Military Command (Koramil). All of these acts occurred either on 7 July 1982, 29 August 1982 or 30 August 1982.

3.2.1.3 The pattern of reported acts of displacement and detention over time

Reported acts of displacement are concentrated in two main time-periods: the third quarter of 1982, when 51.0% (919/1,803) of displacements were reported, and in the fourth-quarter of 1985 when 40.6% (732/1,803) of displacements were reported. This pattern is shown in Figure 83. It was during these two periods that most of the Mauchiga residents who were interned on Ataúro were transferred from Mauchiga to Ataúro and sent back to Mauchiga from Ataúro. At both of these times, individuals were first temporarily transferred to transit locations for short periods of time varying between one day and a few weeks.

Figure 83



3.2.1.4 The pattern of reported acts of displacement and detention over geographic space and by duration of detention

As can be seen in Table 25, the duration of detention periods reported to the documentation project varied widely, lasting from one day to 1,005 days. Around 41.0% (739/1,803) of all reported displacements lasted for ten days or less, 22.1% (399/1,803) of displacement events lasted between 101 and 300 days, and 20.% (368/1,803) between 701 and 900 days. The data collected by the Mauchiga Documentation Project are consistent with the hypothesis that although Mauchiga residents were displaced several times (on average each individual was displaced 3.9 times), around half of these displacements were short-term (of duration less than a week), whereas the other half were substantially longer (ranging from three to 33 months).

Table 25: Distribution of reported duration periods of displacement events of Mauchiga residents, 1982-1985

Duration	Count	% Share
1 day	192	10.7
2-10 days	547	30.3
11-100 days	113	6.3
101-300 days	399	22.1
301-500 days	87	4.8
501-700 days	8	0.4
701-900 days	368	20.4
901-1005 days	87	4.8
Unknown	2	0.1
Total	**1,803**	**100**

Source: Data collected by the Village Secretary of Mauchiga

The Indonesian military held Mauchiga residents in long-term detention in a number of locations, including the island of Ataúro, Dotik, Ainaro Town and Dare. 79.3% (368/464) of the Mauchiga residents documented by the Mauchiga Documentation Project were held in long-term detention on Ataúro.[70] Smaller groups of Mauchiga residents were reported to have been held in long-term detention in Dotik and Ainaro Town: 80 individuals were held for 2.5 years in Dotik starting in November 1982, and seven individuals were held for two years and nine months in Ainaro Town starting in April 1983. Follow-up interviews conducted by the Commission with Mauchiga residents in 2005 in Mauchiga indicated that those individuals who were held in long-term detention in Dotik and Ainaro were held there due to overcrowding on Ataúro.[71]

The broad reported pattern of displacement to and from long-term detention locations for Mauchiga residents was as follows: victims were usually arrested and detained in Mauchiga and then deported to nearby locations where they were kept for anywhere between one day to 10 days. Because these locations were usually transit stops en route to long-term detention centers or back to Mauchiga, periods of detention in them usually occurred in quick succession. The places where people were held for these short periods included Ainaro Town, Bonuk (Hatu Udo, Ainaro), Dare (Hatu Builico, Ainaro), Lesuhati (Hatu Builico, Ainaro), Same (Manufahi) and the Comarca prison in Dili. Others were held in other locations, such as Dotik, Ainaro Town and Dare, before or after their long-term detention periods. Mauchiga residents were also detained for several months in Dare, Dotik, Same, Ainaro Town and Bonuk en-route to or from their long-term internment locations.

[70] The Indonesian military sent 360 of these 368 Mauchiga residents to Ataúro on 30 August 1982. The remaining eight Mauchiga residents were sent to Ataúro on 5 September 1982 after being interrogated for one week in Lesuhati.

[71] CAVR Interviews with Olga da Silva, Abilio dos Santos, Xavier do Amaral and Antonio Pires, Mauchiga, 16 April, 2005.

Table 26: Cross-tabulation of reported duration periods of displacement events of Mauchiga residents by location, 1982-1985

	1 day	2 – 10 days	11-100 days	101-300 days	301-500 days	501-700 days	701-900 days	901-1005 days	Not Known	Total
Ainaro	2	0	16	22	0	0	0	7	0	47
Ataúro	0	0	0	0	0	8	368	0	1	377
Bonuk	0	355	22	0	0	0	0	0	0	377
Comarca	13	0	0	0	0	0	0	0	0	13
Dare	172	0	0	377	80	0	0	0	1	630
Dotik	0	0	6	0	7	0	0	80	0	93
Lesuhati	1	182	0	0	0	0	0	0	0	183
Same	4	10	69	0	0	0	0	0	0	83
Total	**192**	**547**	**113**	**399**	**87**	**8**	**368**	**87**	**2**	**1,803**

The data collected by the Mauchiga Documentation Project are consistent with the hypothesis that the Indonesian military used long-term detention on Ataúro and in Dotik and Ainaro Town of supporters and suspected supporters of the Resistance movement as a strategy to eliminate the social and operational support base of the Resistance in Mauchiga. These data also show that Mauchiga residents experienced a series of displacements and detentions before and after their long-term period of detention on Ataúro, or in Dotik and Ainaro Town.

3.2.2 Reported fatal violations suffered by Mauchiga residents

3.2.2.1 The distribution of reported fatal violations suffered by Mauchiga residents over time and by cause-of-death

The Mauchiga Documentation Project documented 262 unique conflict-related deaths of Mauchiga residents during the Commission's reference period: 44.7% (117/262) of these were killings, and the remaining 55.3% (145/262) were deaths due to illness or hunger. As Figure 84 shows 68.3% (179/262) of these fatal violations occurred during the period of Indonesian military's initial invasion and occupation between 1978 and 1984. Furthermore, the reported patterns of killings and illness/hunger-related deaths of Mauchiga residents are positively correlated.[72] Hence the data documented by the Mauchiga Documentation Project are consistent with the hypothesis that conflict-related killings and illness/hunger deaths in Mauchiga were overwhelmingly concentrated during the Indonesian military's counter-Resistance operations.

[72] The correlation coefficient for these two series is 0.57.

Figure 84

Source: Data collected by the Village Secretary of Mauxiga

While the pattern of killings and illness/hunger-related deaths are correlated over time, there are some notable differences between the two phenomena. Documented killings are concentrated mostly in 1978 and 1982-83 (with 19.7% (23/117) occurring in 1978 and 47.9% (56/117) in 1982-83), whereas 44.8% (65/145) of illness/hunger-related deaths are concentrated in 1983-84 when Mauchiga residents experienced mass deportations from their homes.

3.2.2.2 The pattern of attributed responsibility for reported fatal violations of Mauchiga residents

Of the killings documented by the Mauchiga Documentation Project, 83.8% (98/117) were reported to be the sole responsibility of the Indonesian military, 6.0% (7/117) the sole responsibility of East Timorese auxiliaries and for 10.3% (12/117) institutional perpetrator responsibility was not reported. No reported killings were attributed to the Indonesian military and East Timorese auxiliaries acting together, nor were any killings attributed to individuals associated with the Resistance. This pattern of attributed responsibility appears to distinguish Mauchiga from other parts of Timor-Leste, where a substantial proportion of killings were attributed to Indonesian forces and East Timorese auxiliaries acting together (see above).

Of the documented killings of Mauchiga residents attributed to the Indonesian military, 66.3% (65/98) were suffered by individuals associated with the Resistance and the remaining 33.7% (33/98) by members of the civilian population.

3.2.2.3 The pattern of fatal violations against Mauchiga residents by the victim's political affiliation

The distribution of documented killings and illness/hunger-related deaths varied substantially by the political affiliation of victims. As is shown in Table 27, 64.1% (75/117) of killings were reported to have been committed against individuals formally associated with the Resistance and the remaining 45.9% (42/117) were committed against unarmed civilians. By contrast all but one death due to illness or hunger was reported to have been experienced by an unarmed civilian. This is consistent with the hypothesis that that although killings were mostly targeted against Resistance and clandestine members, the Indonesian military and its associates killed a substantial portion of civilians during its counter-Resistance operations.

Table 27: Distribution of reported fatal violations by political affiliation of victim, 1974-1999

Victim's affiliation	Killings		Illness/Hunger-Related Deaths	
	Count	%	Count	%
Unarmed civilian	42	35.9	144	99.3
Resistance/ Clandestine member	75	64.1	1	0.7
Total	**117**	**100**	**145**	**100**

Source: Data collected by the Village Secretary of Mauchiga

The geographic distribution of reported killings of Mauchiga residents differed from that for deaths from hunger and illness. As Table 28 shows, documented illness/hunger-related deaths were almost uniformly distributed between the sub-districts of Ainaro Town (Ainaro), Alas (Manufahi) and Ataúro (Dili), whereas reported killings were mostly concentrated in Ainaro Town (Ainaro), Alas (Manufahi), and Same (Manufahi).

Table 28: Distribution of reported fatal violations by political affiliation and geographic location, 1974-1999

Sub-district in which violation occurred	Killings		Illness/Hunger related deaths		Total	
	Count	%	Count	%	Count	%
Ainaro, Ainaro	45	38.5	43	29.7	88	33.6
Maubisse, Ainaro	4	3.4	0	0	4	1.5
Bobonaro, Bobonaro	1	0.9	0	0	1	0.4
Alas, Manufahi	31	26.5	47	32.4	78	29.8
Fatuberliu, Manufahi	4	3.4	0	0	4	1.5
Same, Manufahi	31	26.5	0	0	31	11.8
Laleia, Manatuto	1	0.9	0	0	1	0.4
Ataúro, Dili	0	0	55	37.9	55	21
Total	**117**	**100**	**145**	**100**	**262**	**100**

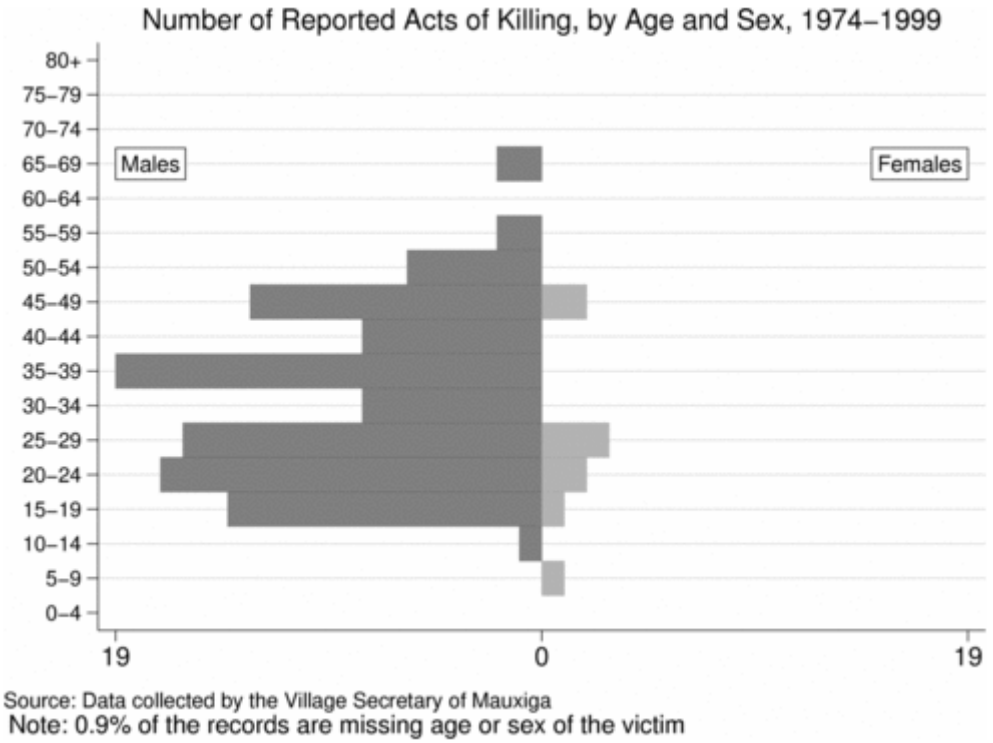
Source: Data collected by the Village Secretary of Mauchiga

3.2.2.4 The pattern of reported fatal violations against Mauchiga residents by the age and sex of the victim

As was the case for killings throughout Timor-Leste, males in Mauchiga suffered the overwhelming majority of killings reported to the Mauchiga Documentation Project. I92.3% (108/117) of reported killings were against males and the balance 7.7% (9/117) were against females). When we move from simple violation counts to population-based violation rates, it can be seen that, on average, relative to their share of the population of Mauchiga Village, the population-based rate at which men were killed was more than 10 times higher than that for women. Ninety-five men per 1,000 were reported to have been killed during the Commission's reference period compared with eight women per 1,000.[73]

As can be seen in Figure 85, 41.0% (48/117) documented killings were against young males between the ages of 15 and 29. This is consistent with the hypothesis that as part of its counter-Resistance strategy the Indonesian military targeted young males of military age.

Figure 85



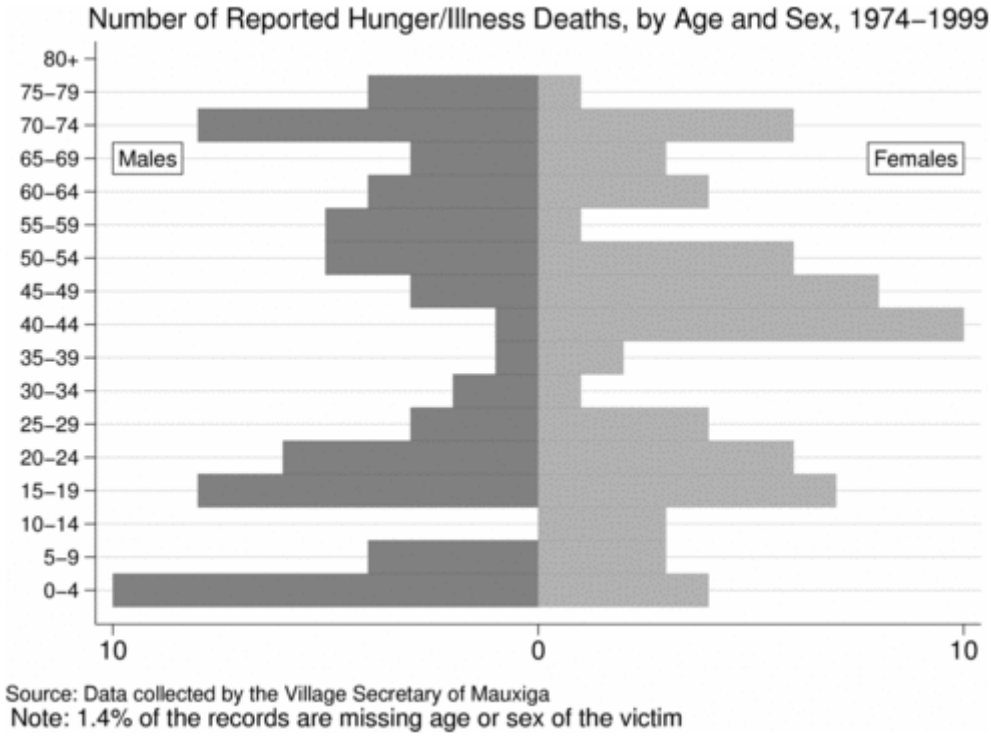
Source: Data collected by the Village Secretary of Mauxiga
Note: 0.9% of the records are missing age or sex of the victim

By contrast documented deaths due to illness and hunger were more evenly distributed across the sexes: 50.3% (73/117) of these were male deaths and 49.7% (72/117) were female deaths. In terms of population share, equal population-based rates of deaths due to illness /hunger were observed for males and females: 64 per 1,000 males in Mauchiga were reported to have died due to hunger/illness during the Commission's reference period, as was also the case for females.[74]

[73] These population-based rates are derived using population figures from the 2001 Timor Leste Suco Survey.
[74] These population-based rates are also derived from population figures in the 2001 Timor Leste Suco Survey.

As can be seen in Figure 86, the residents of Mauchiga who were most frequently reported as suffering deaths due to hunger and illness were young infants and the elderly. This pattern of vulnerability to famine-related deaths of the very young and the elderly is similar to that which was documented by the Commission throughout Timor-Leste.

Figure 86



4. Summary and Conclusion

The Commission for Reception, Truth and Reconciliation collected and utilized a wide array of empirical data sources. In particular, the Commission collected nearly 8,000 narrative statements from Timorese regarding their experiences over 25 years of conflict, conducted a household survey which inquired into mortality and displacement from almost 1,400 households, conducted a census of public graveyards in all 13 districts of Timor-Leste, and developed datasets from information collected from other organization and groups.[75] The purpose of this statistics chapter has been threefold:

1. to present multiple, independent, scientific estimates of the total extent, pattern and trend of mortality and displacement which were experienced during the Commission's reference period,
2. to outline and interpret the descriptive statistics regarding the nature and extent of violations, behavior of perpetrators, and characteristics of victims that were convened from multiple independent data sources, and

[75] External information which was convened into newly established datasets included reports by Amnesty International, narrative interviews collected by the Timorese women's rights organization Fokupers and lists tabulated by village leaders in the Eastern village of Mauxiga.

3. to document the statistical methods employed in reaching the Commission's statistical findings.

In order to develop this depth and breadth of analysis, the Commission and the Human Rights Data Analysis Group jointly developed multiple data projects which involved large-scale data collection, data coding, database representation, record linkage and statistical analysis. The resulting statistical analysis helps to uncover and clarify our social and historical knowledge of political violence in Timor-Leste between 1974 and 1999. However, the Commission's statistical findings are one part of its findings which, like all of its other methodological approaches, needs to be combined and integrated with its qualitative, historical and legal findings.

5. Appendix on Data and Statistical Methods

5.1 Introduction to the Appendix on Data and Statistical Methods

Human rights incidents are complex. An eyewitness or victim may report one or many victims, who may each have suffered one or many violations. Each violation may also involve one or many perpetrators. Hence, the interactions between different persons in thousands of these types of incidents require careful empirical methods of identification and aggregation in order to facilitate valid and reliable quantitative analysis.

To assure the quality of its data, the Commission instituted several processes. This methodological appendix presents the data and methods from which the Commission's statistical results are derived.

The appendix is divided into six main sections. Following the introduction in Section 5.1, Section 5.2 provides detailed descriptions of the different datasets which were used in the Commission's statistical analysis. Section 5.3 describes the data editing, cleaning and name normalization techniques which were applied to the data. Section 5.4 presents the recording accounting tabulations at different stages of the data conversion process. Section 5.5 presents the various deduplication and record linkage techniques which were used to match multiple reports of the same individual victim. Section 5.6 documents the data processing which was used to account for multiple reports of groups of anonymous victims. Finally, Section 5.7 presents the statistical estimation techniques which were used to derive total estimates of the magnitude and pattern of fatal violations and displacements during the Commission's reference period.

5.1.1 Relevance of Empirical Data Analysis to Commission's Mandate

The Human Rights Data Analysis Group (HRDAG) helped the CAVR to collect and analyze human rights violation data relevant to the mandate period of the Commission, 1974-1999.[76] This Appendix explains how the data was organized and processed.

[76] HRDAG is a division of Benetech, Inc. in Palo Alto, CA, USA. HRDAG staff includes statisticians, computer programmers, and record linkage experts. HRDAG team members have worked in large-scale human rights documentation and analysis projects on 5 continents, in over a dozen countries for the past 20 years. HRDAG has worked with official truth commissions in Haiti, South Africa, Guatemala, Peru, Ghana, Sierra Leone; the International Criminal Tribunal for the Former Yugoslavia; and non- governmental human rights groups in El Salvador, Cambodia, Guatemala, Colombia, Afghanistan, Sri Lanka, and Iran. For more information see http://www.hrdag.org.

The CAVR required an information management system to manage and structure the data needed to answer the issues outlined in its mandate. Specifically, the CAVR's information management system had to supply information about past human rights violations which would subsequently allow for:

1. Descriptive statistical analyses of general patterns and trends of violations in order to describe the "nature" of human rights violations (i.e. nature in terms of the types of violations which were committed)[77]
2. Statistical projections of total violations to establish the "extent" of human rights violations (i.e. extent in terms of the total number of violations which were committed)[78]
3. Statistical hypothesis testing of the regularity of certain violations in order to investigate whether certain violation patterns constituted "a systematic pattern of abuse"[79]
4. Case-level analysis by basic filing and searching of the database in order to describe the "antecedents, circumstances, factors, context, motives and perspectives" that led to large-scale violations [80]
5. Structured quantitative analysis and hypothesis tests in order to investigate whether "human rights violations were the result of deliberate planning, policy or authorization" on the part of specific parties to the conflict,[81]

and

6. Formal explanations of scientific and statistical methodologies employed in the final report's appendix in order to demonstrate that CAVR findings are based on "factual and objective information and evidence collected or received by it or placed at its disposal"[82]

The CAVR was particularly sympathetic that a large proportion of victims and their families have lived in silence, fear, and isolation about violations which people suffered as far back as 1974. Therefore the Commission's data collection and information management choices had to both produce reliable historical data and promote public participation in the truth-seeking process.

5.2 Data Sources

This section details the three primary statistical databases the CAVR established to complete quantitative analysis of past human rights violations and promote reconciliation. The Human Rights Violations Database (HRVD) was a collection of narrative statements from deponents to the Commission's statement-taking process, qualitative reports from Amnesty International (AI) and data collected by Fokupers, a local Timorese NGO. The Retrospective Mortality Survey (RMS) was a random-sample household survey used to measure displacement and mortality during the CAVR's mandate period. The Graveyard Census Database (GCD), was a comprehensive census of public graveyards in each of Timor-'s thirteen districts.

---

[77] UNTAET Regulation 2001/18 Section 13.1(a)(i)
[78] UNTAET Regulation 2001/18 Section 13.1(a)(i)
[79] UNTAET Regulation 2001/18 Section 13.1(a)(i)
[80] UNTAET Regulation 2001/18 Section 13.1(a)(ii)
[81] UNTAET Regulation 2001/18 Section 13.1(a)(iv)
[82] UNTAET Regulation 2001/18 Section 13.1(d)

The combined data from all three of the Commission's data streams - the HRVD, the RMS, and the GCD - were used to make independent demographic estimates of the total extent, pattern, trend, and levels of responsibility for past fatal violations in Timor-Leste.

5.2.1 Human Rights Violations Database (HRVD)

The following sections describe the three documentation projects which were convened to form the Commission's Human Rights Violations Database. The process of transforming qualitative information from these documentation projects into statistical data is also presented. Finally, the recording accounting from the three documentation projects is presented.

5.2.1.1 CAVR's statement-taking process

In February 2003, the Commission began collecting narrative statements from individuals in all thirteen districts of Timor-Leste and from Timorese then living in West Timor. These statements were the basis of the HRVD. The CAVR established offices in each of the country's thirteen districts as part of the CAVR's socialization and reconciliation process. A total of 7,669 narrative statements were collected documenting reported human rights violations. These narratives provided extensive information on both fatal and non-fatal violations during the reference period.

In order to analyze this qualitative information statistically, it was coded into a FoxPro database using the design standards of the "Who Did What To Whom" data model.[83] Although these data provide many useful insights, the CAVR statement-taking process that generated them did not employ a probability-based random sample. Rather, the CAVR accepted statements from those willing to volunteer the information they could recall. As a result, the narrative data, in isolation, cannot be assumed to be statistically representative of the overall extent and pattern of violations in Timor-Leste.

5.2.1.2 Amnesty International

The London-based human rights advocacy group Amnesty International reported on the Timorese human rights situation during the Commission mandate period mostly by way of information gathering through underground networks in Timor-Leste and through its contacts within the Timorese diaspora in Australia and Portugal.

The Commission received 322 reports and documents from Amnesty International, which were compiled between 1985 and 1999.[84]

Amnesty International's qualitative reports and "Urgent Actions" were coded and entered into the Commission's human rights violations database using the same methods and standards as were

[83] Ball "Who Did What To Whom Handbook" and "Ball et al: HR Database Design Methods"

[84] The Commission was unable to locate the following Amnesty International Reports:
ASA 21/12/83 UA 212/83 21 September
ASA 21/16/85 Disappearances
ASA 21/44/85 Unfair Trials and Possible Torture in East Timor
ASA 21/22/87 Statement on ET by AI to the UN Special Committee on Decolonisation
ASA 21/23/87 ET: Releases of Political Prisoners
ASA 21/14/91 AI statement to UN Special Committee on Decolonisation - Appendix I and II
ASA 21/24/91 East Timor: After the massacre - Appendix 1
As a result, the Commission's statistical analysis of violations in Timor-Leste reported by Amnesty International does not include relevant acts and incidents covered in these reports.

used for the statements which were collected by the CAVR. The information collected from Amnesty International describes the general human rights situation in Timor-Leste, as it was observed by the international human rights community at the time.

5.2.1.3 Fokupers

Forum Komunikasi Untuk Perempuan Loro Sae (Fokupers) a local human rights NGO, constructed a violations database after the referendum-related violence in 1999.[85] The Fokupers database is constructed from open-ended interviews conducted by Fokupers' staff with local Timorese women. Originally, the main purpose of the interview was linked to the counseling work conducted by Fokupers. However, the objectives were extended to include documentation for investigation purposes by competent legal authorities, such as the UN's Serious Crimes Unit. These narrative statements were taken in the Tetum language.

Fokupers constructed its database to facilitate the publication of a report on violence against women. Their original database was centered around representing the biographical data of victims, the narrative events that were described, identifying the violations which occurred and perpetrators involved. In July 2004, Fokupers submitted this data to the Commission on the condition that personal identifiers of deponents, victims, or family members in the database would not be identified in the CAVR report. CAVR staff recoded the data, based on the Commission's standardized definitions and coding scheme, so that these data could be analyzed in parallel with the CAVR's human rights violations database.

5.2.1.4 Coding the Qualitative Sources (CAVR Narrative Statements, Amnesty International & Fokupers)

Data coding is the process of transforming unstructured narrative information on violations, victims, and perpetrators into a countable set of data elements, without discarding important information or misrepresenting the collected information.

In October 2003, the Data Processing Staff reviewed the coding and data entry process in order to identify systematic errors and inconsistencies in the coding and data entry process. At the time, 2,473 statements had been coded and entered into the Commission's database. A random sample of 15% of statements (i.e., 371 statements) in the database was taken, stratified on the district in which the statement was taken.

Each statement was reviewed by a coder: the coder re-coded the statement without looking at how it had been coded originally. Then the results of the two codings were compared and errors in the original coding were identified, noted and then changed. In addition, the coder would also review the database entry for this statement and identify and note any data entry errors and correct them.

Within the 371 reviewed statements, 416 coding errors were identified. 58% (241/416) of these errors were violation coding errors, 12% (49/416) errors associated with coding of the victim's affiliation, 10% (42/416) with the level of location specificity coded and 9% (36/416) were associated with the coding of the institutional perpetrator responsibility. Of the identified 416

[85] Fokupers, the East Timorese Women's Communication Forum, was founded in 1997 and focuses on support to victims of political violence through counseling programs and other forms of assistance to women victims of violations, including ex political prisoners, war widows, and wives of political prisoners. Its mandate also includes promoting women's human rights among the local population, especially East Timorese women.

errors, 70% (291/416) of these coding errors were errors of non-identification (i.e., where the act was not identified as a violation or the person or location was not identified by the coder). Another 17% (71/416) of the coding errors resulted from the coder including the act when what was described in the narrative did not met the definitions and boundary conditions of the Commission's controlled vocabulary. Finally, 13% (54/416) of the coding errors were the result of misclassification of an act into the incorrect violation category.

As a result of this coding review, the Data Processing team undertook three initiatives to minimize these errors in the future: (1) a number of revisions were made to the Commission's controlled vocabulary, (2) a training workshop in which the results of the review were presented to the coding team and extra training provided in the necessary areas, and (3) the implementation of regular group coding exercises where coders coded the same statements and reviewed the consistency of their coding decisions using both qualitative reviews and quantitative inter-rater reliability (IRR)[86] measures.

The main types of revisions which were made to the Commission's controlled vocabulary were:

- a reduction in the number of violation categories to a more manageable list
- refinement of boundary conditions for conceptually similar violation categories (such as torture and ill-treatment)
- refocusing the controlled vocabulary to focus on measurement of violations only, not both the measurement of violations and the physical and psychological impact of these violations
- simplifying the definitions of violation categories and ensuring the syntax of the definition is more consistent with the specificity of information collected in the statements (e.g., technical legal terms were reworded into common language or eliminated, as they did not fit the historical reality being measured)
- revision to the institutional actors list: both simplification of the list and hierarchical structuring of the institutions to reflect their structural relationships with each other.

5.2.1.5 HRVD Data Collection Results

The HRVD's three combined data sources produced a database with records as shown below in Table 29. These records represented individual and group victims, both of which suffered fatal and non-fatal violations. Table 29 shows the breakdown of the number of records collected in each database. Note that these numbers represent the data totals before cleaning where invalid and duplicate records were removed from the databases.

[86] Inter-rater reliability is the extent to which two or more coders agree. Inter-rater reliability addresses the consistency of the implementation of a coding system. For a detailed discussion of how statistical IRR measurement can be applied to human rights violation coding, refer to Romesh Silva ´On the Maintenance and Measurement of Inter-Rater Reliability when Documenting Large-Scale Human Rights Violations¡ Proceedings of the Joint Statistical Meetings of the American Statistical Association, the International Biometric Society (ENAR and WNAR), the Institute of Mathematical Statistics, and the Statistical Society of Canada. August, 2002.

Table 29: Recording Accounting Matrix for the Human Rights Violations Database

Data Source	Statement Count	Individual Count	Fatal Violations	Non-Fatal Violations
CAVR statements	7,779	38,812	6,778	31,595
Amnesty International	267	547	122	631
Fokupers	423	4,888	376	3,983
Totals	**8,469**	**44,247**	**7,276**	**36,209**

Groups are records of unnamed victims that identify two or more victims. Some victims suffered multiple non-fatal violations, others suffered non-fatal violations and a fatal violation, and others suffered only a fatal violation. Consequently, violation totals do not sum to the victim count.

5.2.2 Retrospective Mortality Survey (RMS)

The CAVR undertook a Retrospective Mortality Survey to provide a probability-based estimate of displacement and deaths. This survey drew a stratified random sample of households, and used a structured questionnaire to collect information about deaths in the family and displacement events during the Commission's reference period. The survey enabled statistical estimates of the extent of natural mortality, famine related deaths, conflict-related deaths, and migration.

5.2.2.1 Statistical Sampling used in the RMS

The RMS sample was based on a two-stage sample design. The first stage was a sample of all 2,336 aldeias in Timor-Leste, and the second stage was a sample of households within the selected aldeais.[87]

The population of households was stratified along the following variables: urban/rural, district location, and elevation.[88] Implicit stratification methods were used so that the list of aldeias was sorted by the following ranked variables: urbanicity, district, and altitude, and a systematic random sample picked aldeais across each of the stratification variables.[89] A cumulative measure of size variable is created and a sampling interval is calculated as the number of clusters (144)

[87] An aldeia is the smallest administrative unit in Timor-Leste. In general, an aldeia is a settlement of group of homes in a small local area. Usually, a suco (village) is made up of three or four aldeias, and groups of sucos make up a sub-district which is an administrative subset of a district. There are 13 districts, 64 sub-districts, 498 sucos, and 2,336 aldeias in Timor-Leste.

[88] Stratification is the process of grouping members of the population into relatively homogeneous subgroups before sampling. The strata need to be mutually exclusive such that every element in the population may be assigned to only one stratum. The strata should also be collectively exhaustive, in that no population element can be excluded. Random sampling is then applied within each stratum. Stratified random sampling often improves the representativeness of the sample by reducing sampling error.

[89] We used a method known as Probability Proportional to Size (in this case "size" refers to the number of households and not population, although the two are obviously correlated), a common design in surveys of this kind.

divided by the total measure of size (180,015), which equals 1250.1. A random number between 1 and 1250.1 was generated (397.235) and the aldeia with a cumulative measure of size above that number was selected in the sample. 1250.1 was added repeatedly to the initial randomly generated number and aldeias were selected throughout the list in the same fashion.

The decision to draw a fixed number of 20 households, instead of something proportional to the size of the aldeia or some other allocation method, is primarily one of operational considerations. Selecting a fixed number of households per aldeia is one way of retaining control of the overall sample size and of having an approximately uniform distribution of workload among interviewers.

The CAVR considered the feasibility of incorporating Timorese respondents still displaced in West Timor into the reference population.[90] However, a number of security, operational, and data quality concerns in West Timorese refugee camps made survey implementation in West Timor difficult. Therefore, the reference population that was sampled by the Commission consisted of all households within the thirteen districts of Timor-Leste.

It was not optimal, for both statistical and operational reasons, to allow aldeias with fewer than 20 households to be sampled. Therefore, small aldeias were combined with nearby aldeias (which were not necessarily adjacent), before sampling took place, so that the estimated number of households in a cluster (defined as an aldeia or group of aldeias) was at least 40, to reduce the chance that a sample cluster had fewer than 20 households. Although upon the arrival of our field team it was still possible due to have fewer than 20 households due to the inaccuracy of the frame. That is, the number of households reported in the 1990 census could have been either originally inaccurate, or it could have changed by the time the field team arrived in 2003. Therefore, the 144 sampled aldeia cluster actually contains 165 aldeias. Operationally, this means that in these clusters, interviewers had to draw a random sample of 20 households from among the combined total number of households in the cluster.

5.2.2.2 Questionnaire Design & Development for the Retrospective Mortality Survey

The questionnaire of the RMS was designed to fulfill the following objectives:

- to produce estimates of total mortality in Timor-Leste between 1974 and 1999, using both survey-based estimation techniques and Multiple Systems Estimation techniques, and
- to develop survey-based analysis that estimate and describe the complicated displacement movements within Timor-Leste throughout the Commission's mandate period.

As a result, the questionnaire was organized into the following modules:

- household register
- head of household displacement register
- adult female birth history
- adult male/female sibling history

[90] Section 3.3 Regulation 2001/10 "The Commission may conduct all such activities that are consistent with the fulfillment of its mandate within the present Regulation."

- adult male/female parental history
- general human rights violation section

The questionnaire[91] was reviewed by three human rights statisticians[92] external to the Commission and several subject specialists at the Commission. Through this review process, improvements were made to the layout and design of the questionnaire, and a number of terminology issues in Indonesian and Tetum languages were identified.

With local staff, Jana Asher conducted eight cognitive interviews during the questionnaire development phase. Cognitive interviewing explores the cognitive processes of the respondent. It seeks to identify difficulties and possible solutions to challenges faced by respondents in (i) comprehension of the question, (ii) retrieval from memory of relevant information, (iii) decision processes, and (iv) response processes.[93] A total of eight subjects - four in laboratory conditions and four in the field - participated in the cognitive interviewing. Significant insight was gained from the probing on respondent's date recall. In particular, cognitive processes and responses about time and date related questions indicated that often, when a respondent answered "Don't Know", they may just not know the exact date according to the Gregorian calendar. However, their responses indicated that sometimes notions of time were easier to recall in relation to major events, or agricultural or seasonal variation.

From the cognitive interviewing process, structured date probes were developed which asked the respondent to narrow event-dates into a six-month window which could be defined by major events such as holidays, or environmental/physical indicators (height of corn or other crops, rainy season or dry season). The cognitive interviewing process also indicated that temporal concepts such as beginning, middle, and end were not understood by respondents, so further narrowing of the time window was not possible.

During the cognitive and field test interviews, respondents often simply answered "Don't Know" or "into the mountains" as the place to which they were displaced. As a result of the cognitive interviewing, a careful set of probes was created to elicit more detailed descriptions of the places where people were displaced.

After peer-review and the cognitive interviewing process, the finalized questionnaire was then translated and back-translated into both Indonesian and Tetum. The questionnaire was then field tested for 5 days in aldeias within Dili which were not part of the sample. As a result of this field test, a few further question-sequencing, grammatical, and syntactical improvements were made.

5.2.2.3 Survey Implementation & Fieldwork

Within each sampled household, the head of household responded to both the household register (in which all residents of the household were logged) and the displacement section. An adult female was then randomly selected from the female adult population of the household to answer the adult female birth history module.

---

[91] Reference Section 5.8 for copy of survey

[92] Fritz Scheuren, Ph.D., President of the American Statistical Association, consultant to HRDAG on projects for Kosovo, Guatemala, and Peru; William Seltzer, Fordham University, and Jana Asher, co-author of HRDAG reports in Kosovo, Sierra Leone, and Peru.

[93] Tourangeau 1984.

Before leaving each aldeia, all questionnaires were checked by field supervisors to identify and correct any mistakes or inconsistencies in the completed questionnaires. Two field coordinators accompanied the team of 22 survey enumerators into the field.

Twelve aldeias which were included in the sample, were not able to be visited by the enumeration team. The team was unable to conduct interviews in these twelve aldeias due to security concerns at the time. Table 30 lists the twelve aldeias that were not enumerated.

Table 30: Aldeias which were in the RMS Sampling Frame which were not visited by the RMS Enumeration Team

District	Sub-district	Suco	Aldeia
Alieu	Remexio	Sucu Liu Rai	Coto Mori
Baucau	Fatu Maca	Samalari	Osso Luga
Baucau	Laga	Sama Lari	Soru Gua
Bobonaro	Atabae	Atabae	Heleso
Bobonaro	Bobonaro	Tapo	Tapo
Cova Lima	Fohorem	Datorua	Fatulidun
Lautem	Iliomar	Ailebere	Heitali
Lautem	Lospalos	Fuiluro	Kuluhun
Liquiça	Bazartete	Fahilebo	Fatu Neso
Oecussi	Passabe	Abani	Na Nos
Viqueque	Ossu	Uai Bobo	Sogau
Viqueque	Uato Lari	Mata Hoi	Loko Loko

Furthermore, in some aldeias less than ten households were enumerated resulting in some additional non-response. Overall, of the 1,440 households in the sampling frame, there was a 3.1% (44/1440) non-response rate. Given the low non-response-rate, no explicit statistical imputation was performed to control for non-response in the survey.

5.2.3 Graveyard Census Database (GCD)

In order to develop baseline mortality data for Timor-Leste, the Commission undertook a census of public graveyards in Timor-Leste's thirteen districts. Through this process, available information about names, dates of birth, dates of death, and religion was collected. Gravestones that lacked such information were also enumerated and their size was noted.[94] By collecting this

[94] Size of an unmarked gravestone can be used as a proxy indicator of whether the deceased was a child or adult.

information, the Commission created a de facto vital registration system for the Timorese population. That is, the GCD created a baseline listing of some, perhaps even most deaths, which could be used for mortality analysis beyond this project.

5.2.3.1 GCD Data Collection

To facilitate the CAVR's census of public graveyards in the country, a list of all known public graveyards in Timor-Leste was enumerated by CAVR field staff in consultation with village-level officials at the suco level, and where possible, the aldeia level. A "public graveyard" in this study was defined as a location which is reserved exclusively for burial of deceased persons. This definition includes communal burial sites which are established on public land or land owned by a religious institution. However, it excludes family graves located on private property.

The GCD data was collected by two separate data collection teams. The first team collected 128,751 records from 803 cemeteries, which were entered into an series of Excel spreadsheets. The first team covered portions of all thirteen districts, but only Dili was covered completely. A second team went into all districts, except Dili, to finish the graveyard survey. They collected 153,057 additional records from 1,779 cemeteries. The second team used a FoxPro database for their data entry.

The CAVR enumeration teams documented all gravestones within public graveyards – both marked and unmarked. A marked grave was defined as having a physical structure which memorialized a person's life, with legible markings in English, Indonesian, Tetum, or Portuguese.[95] On all enumerated marked gravestones, the following information was coded if on the gravestone: full-name, date-of-birth, and date-of-death. Unmarked gravestones were typically small simple crosses or other burial markers, without name or date information for the deceased. Enumerators were asked to note information about the religion, type-of-material and grave size, if it was discernible from the gravestone, for both marked and unmarked gravestones.

5.3 Methodological Description of Data Editing, Cleaning & Name Normalization Techniques

Each of the three databases used by the CAVR required data editing, cleaning, and name normalizing techniques in order for the data to be compared and linked between databases. Several months were spent reviewing the data for obvious typographical or spelling errors, and a random sample review was conducted to ensure data accuracy. Technical problems occurred in converting data from one database structure to another, and these were also identified and corrected.

5.3.1 Database Cleaning and Editing

The Data Processing staff carried out a complete check (and correction where required) of all HRVD records with:

- missing district/sub-district information
- implausible violation date information (e.g., day = 42, month =13)
- records where the violation occurred before the victim's birth date

[95] Due to a lack of resources, we were unable to enumerate Chinese graveyards.

- records where the violation occurred after the victim's death date
- statements where the deponent was coded as a victim of a fatal violation
- records where the victim's age was coded as 0 or as a negative number
- records where the victim's age was coded as greater than 75
- records where there was no violation code recorded
- records where there was no victim recorded for a coded violation
- records where there was no (individual/institutional) perpetrator assigned to a coded violation

In addition to the complete quick-checks described above, the coding team also did checks of a simple random sample of records of fatal violations, detentions, tortures, ill-treatments, forced recruitment, sexually-based violations and displacements. The objective of the quick checks was to identify whether there were any systematic errors in assigning affiliations of victims and institutional perpetrator responsibility. One major inconsistency was identified - namely where victim affiliation was not assigned to all victims of a violation(s) which happened in the same act or acts closely linked in time. These records were identified, and appropriate rules were applied to correctly assign victim affiliation across violations in the same act or proximate acts for the same actor.

5.3.2 Date Editing & Cleaning

Records that had obvious errors, such as dates of birth, violation, or death that were subsequent to the current date were examined and corrected. This was especially common in the GCD database where the grave markers were so small that full four-year dates could not be written out. The data entry system defaulted the two year dates, which should have been in the 1900's, as the 2000's. Enumerators from different teams sometimes used different date coding standards. Some used the European standard DD-MM-YYYY, the US standard MM-DD-YYYY", YYYY-MM-DD format, or variations of these using a two-digit year. Furthermore, sometimes different separators were used between years, months and days – including "/", ".", and "-". As a result, all date formats across all three datasets were mapped to the following standardized format; YYYYMMDD.

If the DOB was after the DOD, the dates were swapped. Two types of errors which caused dates with months greater than 12 or days greater the 31 were also identified and examined. We discerned that some errors were caused by variations of the spreadsheet date format settings on the data entry computers.

Other errors were obviously typographical. Records from the HRVD and the RMS were corrected by reviewing the original paper material and applying corrections to the database. For the GCD database, there was not enough time to hand review the source, so if the error was not easily corrected, the values in that part of the date field (month or day) were left blank.

5.3.3 Age Editing & Cleaning

Age data was examined for possible typographical errors, for example, people over the age of 100. The sources for these records were reviewed to verify the data and corrections made as

necessary. Where DOB and DOD information was known, the age was derived. The GCD age value was calculated and a new field generated to facilitate easier matching.

5.3.4 Violation & Relationship Codes Editing & Cleaning

Reviews were conducted of the violation codes and relationship codes within the HRVD and RMS identified codes that were not valid or conflicted with other data within an individual record (e.g., a female being coded as a father). The paper source files for these records were reviewed and the corrections made to the database.

5.3.5 Geographic Location Code Editing & Cleaning

The geographic location data collected for the RMS and HRVD databases was coded to the Timorese geocode standards established by the government and approved for use by the CAVR. Locations were divided into four administrative levels - District, Sub-district, Suco, and Aldeia. For those locations that were outside of Timor-Leste, codes for West Timor and Java were created and when the location was not known, they were marked to a separate code for unknown. Each cemetery was given a unique code, called an "id," in order to differentiate between cemeteries in the same geographic area.

The GCD was not collected with the Timorese geographic code standard, so we translated it to the standard codes.

5.3.6 GCD De-Duplication of Cemeteries and Graves

Several factors lead to duplicate records of graves and graveyards in the GCD database.

1. Different data collection teams unknowingly covered the same cemetery. Many cemeteries did not have posted names, making it hard to identify duplicated records strictly by cemetery name.
2. The exact suco and aldeia location was often hard to determine in some rural areas. Even if the cemetery had the same name, it might be coded to a different geographic location. Additionally, many cemeteries shared the same name (Santa Cruz being the most common name) which meant that cemetery name alone was not enough to determine duplicate cemeteries coded to different geographic codes.
3. Many of the cemeteries in Timor-Leste were not organized linearly. This sometimes led to the team of enumerators crossing over the same gravestone, recording it more than once.
4. Because of the massive amount of paper files required to gather all this data, it was possible that there were data entry duplications.

It was possible to find linkages between cemetery id's by examining the names of the deceased, cemetery locations, cemetery names, and complete dates of birth and dates of death matched.[96] When rows of duplicates were found, one of the cemeteries was dropped from the dataset used for analysis. While it is common for people to have the same forename and surname, and potentially date of death, it is highly unlikely that they would have both the same dates of birth

[96] A complete record is defined as having day, month and year for both DOB and DOD.

and death. Therefore, any records that had the same forename, surname, date of birth, and date of death were considered duplicates, and only one record was kept in the database for analysis.

The goal of the GCD de-duplication process was to ensure that the deceased were counted only once. It was initially thought that during the forced displacements people may have initially been buried where they died, with the body later retrieved by the family and interred at a cemetery in their home aldeia. It was also thought that if the body was not recovered, that a memorial marker in the local cemetery would be erected. While this may have occurred, careful review of the data did not reveal reburial or post-hoc marking with a memorial stone to be common practice. Furthermore, when the bodies were recovered, the first marker would likely have been removed or relocated with the body, thus preventing over-counting. People who were never buried, or who not were buried in public cemeteries, fall outside of the GCD. In order to account for the deaths that are missing from the HRVD testimonies, the RMS interviews, and the GCD grave data, we conducted multiple-system estimation of the total deaths. This analysis is described below in Section 5.7 of this Appendix.

5.3.7 Name Cleaning Processes

The names of persons in the CAVR data needed to be addressed in two ways. First, the names needed to be parsed into three categories – first, middle/nick, and last – names. Once this was complete, name canonicalization was required to facilitate record linkage. Canonicalization is a process of reducing each name to the simplest and most significant form possible, without loss of generality.

Person names contained a significant amount of variation in the spellings, apportionment to the three name fields, and in punctuation. Name variation has many causes. In open-ended narrative statements, such as the HRVD, the deponent may be a close relative, friend, neighbor or distant acquaintance of the victim, and he or she may or may not know how to spell the names of the reported victim. Transcription by the statement-taker may involve application of additional spelling and punctuation rules and even incorporate spelling errors. Similarly, spelling and punctuation transformations may take place at the data coding and data-entry stages.

5.3.7.1 Name Parsing

To address the significant variation in how names were apportioned to the three name fields; first, last, middle/nickname, the names were parsed according to strict rules. The Commission decided to divide the names using the 'first' first name for first, and the 'last' last name as last, and all other names placed into the middle/nickname field. Additionally, the prepositions (e.g., de, da, do, dos) were dropped from the name fields as their use was inconsistent in the data.

For example, the Portuguese name Maria Louise da Costa da Silva may be been entered into the database as shown in Table 31.

Table 31: Hypothetical ways in which a given name (e.g. "Maria Louise da Costa da Silva") might be initially represented in the database

First	Middle/Nickname	Last
MARIA LOUISE		DA COSTA DA SILVA
MARIA	LOUISE	DA COSTA DA SILVA
MARIA LOUISE	DA COSTA	DA SILVA
MARIA	LOUISE DA COSTA	DA SILVA
MARIA LOUISE		SILVA

The name parsing process would have standardized these names so that the first name was Maria while the last name would simply be Silva. All other names, less the prepositions, were moved into the middle/nick fields.

The animist name MauBere may have been entered in the forms shown in Table 32.

Table 32: Hypothetical ways in which a given Animist name (e.g. "MauBere") might be initially represented in the database

First	Middle/Nickname	Last
MAU BERE		
MAUBERE		
MAU		BERE
		MAUBERE

The name parsing in this case would place Mau in the first name field and Bere in the last name field.

5.3.7.2 Name Canonicalization

Name canonicalization was applied to the first and last name fields of the records after parsing to facilitate easier matching, especially the automated algorithms for record linkage. Spelling variants for names were distilled into a single representative form for each name. For example, the following spelling variations were canonicalized to AGUSTINO:

AGUSTINUHO	AGUSTINO	AGUSTIMHO	AGUASTINHO
AAGUSTINO	AGUSTINU	AGSSTINHO	ANTGOSTINHO
AGUSTIO	AGUSTONIO	AGSTINHO	AGUSTINHU
AGUSTINUS	AGUSRINO	AUGUSTINO	AGOTINHO
AUGUSTINHO	AGUSTINHO	AGOSTINHO	AGOSTINO

The animist names were harder to canonicalize as they were generally four or five characters long and some records that appeared to be spelling variations were in fact distinctly different names. Conservative canonicalization was applied to the animist names and then tested with sample linkage of animist records looking at date, age and place information to determine additional canonicals to apply.

After several passes over the names to canonicalize them, a new field was generated with the name spelled out in reverse order. Then, by sorting on this new field, we were able to find additional names to be canonicalized to a single form as beginning letters could vary depending on pronunciation, but the ending syllable was likely the same. This process proved to be very helpful in finding additional canonicals.

There were Chinese, Indonesian (Muslim), and Anglo-Saxon names in the databases, as well as Portuguese names and animist names. The relatively few numbers of Chinese, Indonesian and Anglo-Saxon names did not require special handling. Timorese staff, in Timor, identified whether names were animist for the application of matching rules and algorithms, because animist names are not always sex-specific.

The HRVD and RMS databases are smaller then the GCD, so we canonicalized them first. We then applied the lists of name canonicals to the GCD. The resulting names were then reviewed to identify additional canonicals.

During the canonicalization process, some letters in names were found to be interchangeable with each other, most commonly with the Portuguese names. The letters S, J, G, and Z were often interchanged with each other in names. Also, the letters V, U, W, and B were often interchangeable. Less frequently, the letters H and E were interchanged. An example of interchangeables would be for the name Virginia, which could be spelled with a B or V. For example, spelling variations found for the canonical VIRGINIA included BIRGINIA, BERGINA.

Names that began with these letters were compared to each other to assist in the canonicalization process. Where names had more than one interchangeable or the interchangeable letter was in the middle or end of a name, it was very difficult to find potential canonicals. Therefore, a program was written that generated a list of names where combinations of interchangeable letters matched another canonical name. The record linkage expert reviewed these combinations to determine if they should be canonicalized or were distinctly unique names. Where there were additional canonicals due to interchangeables, they preferred letter for the canonical was S (for S, J, G, and Z), V (for V, U, W, B), and H (for H and E).

Additionally, in the canonical process, it was noted that ANJU and ANJO were often cited as the first name or the only name for a record. Angu is the Tetum word for infant and was found often in the GCD records when a child died before being baptized and therefore was not given a Christian name. Records with ANJU and a last name were used for the matching process because there was some identifying data, but records with only ANJU were too ambiguous to make reasonable judgments for matching.

5.3.7.3 Sex & Ethnicity Coding

During the canonicalization process, the Portuguese first names were reviewed with the frequency of the sex codings male, female and unknown displayed.[97] Sex codings that were obviously incorrect were corrected. As with most Latin names, those that end with A generally are female and those ending with O (or U) are usually male. Where first names ended in letters other than A, O or U, the frequency between male codings and female codings were examined and when the disparity was great, indicating that a few records were miscoded during data entry, corrections were made to the database.

5.4 Data Conversion

In order to expedite all the data processing steps associated with matching of duplicated records, each dataset was transferred from its original FoxPro or Excel database platform, to our Analyzer database platform.[98] The FoxPro database schema was first duplicated in PostgreSQL for importing into Analyzer. The relational database structures for the HRVD and RMS data were maintained in Analyzer.

Table 33 shows the total number of records from each dataset that were imported into Analyzer. Note that these totals reflect data cleaning changes which resulted in the dropping of duplicate and invalid records.

Table 33: Total Record Count by Database
Pre & Post Data Cleaning

Database	Pre-Clean	Post Clean
HRVD	41,546	37,651
RMS	4,883	4,619
GCD	195,468	149,087

5.5 Record Linkage Overview

Individuals reported in the HRVD and the RMS are sometimes reported multiple times, by different deponents and may also appear as records in the GCD. To ensure the statistical analysis controlled for duplicate reports of the same person, the data required record linkage, also known as matching. Matching was applied to two general categories of violations for study -- fatal and non-fatal violations. Fatal violations included civilian killings, deaths due to deprivation, disappearances, and combatant deaths. Non-fatal violation categories included; attempted civilian killing, detention, torture, rape, sexual slavery, sexual violence, ill-treatment, displacement, forced marriage, impediments to reproductive rights, unfair trial, destruction of homes, destruction of livestock, extortion, threats, forced recruitment, and forced labor.

[97] Frequency is a count of the instances a name or code appears in a particular data field. Values with very low frequencies can reveal potential errors or misspellings in the data.
[98] 'Analyzer' is a free and open source application used to collect, maintain, and analyze information about large-scale human rights violations. For more information about Analyzer, see HDRAG website at http://www.hrdag.org/resources/data_software.shtml.

There were two types of matching done for the purposes of statistical estimates; intra- and inter-system matching. Intra-system matching links records that identify the same person within a single dataset, and each record can match to zero, one, or many other records within that dataset. Inter-system matching joins two or more lists of unique records from different data sources together so that a multiple systems estimation of the violations can be applied. Records matching during inter-system matching can match only to zero or one other record in each of the other datasets.

Due to the complexity of inter-system matching and time constraints for the work, the non-fatal data in the HRVD and RMS only had intra-system matching performed for descriptive statistics. The fatal data, which included the GCD data, was both intra- and inter-system matched as the basis for multiple systems estimate calculations. Matching was done using three methods: hand-matching, computer-generated matching, and computer-assisted matching. Each of these methods may involve more than one pass.[99]

5.5.1 Matching Rules

Each individual record was compared to all other records in each dataset for possible matches and was deemed a match when a significant number of the field values match *exactly*, were in *close proximity*, or did *not conflict*. The fields used for matching were: first_name, last_name, age, sex, DOB, DOD, place_of_birth (POB), and place_of_death (POD). The middle/nickname and interview_location fields were also available for clarification purposes, but were not fields available in all three datasets, and was often sparse where it was available. While not part of the matching rules, this data was taken into consideration by the record linkage expert. However, it was not used in any computerized auto-matching.

The matching decisions used for the CAVR data tended to over-match records.[100] Over-matching reduces the number of unique records and therefore will tend to lower the estimates. Over-matching is preferred in cases where there is uncertainty that a match is accurate, to produce conservative estimates.

5.5.1.1 Matching Names

The first and last name fields were not always complete; some had initials or were missing either the first or last name. Attempts were made to match every record even when it was incomplete, but for fatal matching, records with neither first or last names or had initials only, were dropped from matching as there was not enough data to make reliable judgments. For non-fatal matching, attempts were made to match violations with DOB, DOD, and death location information to other records with the same values in those fields, even when there was no name or the record only had initials. Records with less complete name data relied more heavily on perfect dates and places to be matched to other records. Many people could have died on the same day in the same place, and knowing which of those people to match an incomplete name to is difficult and unreliable.

[99] A pass is a review of all the data in a dataset based on sort order or algorithm, to look for matches.
[100] Over-matching means that linkages are made between records that might not in fact be duplicates.

5.5.1.2 Matching Sex & Ethnicity

Where the sex of the victim was known, it was only potentially matchable to records of the same sex or those with unknown sex. Records where sex was marked Unknown were matchable to records coded Male and Female, but within a matched group, the sex codes could not conflict with other records in that group.

5.5.1.3 Matching Locations

Geographic location codes used for the CAVR data were divided into four levels: district, sub-district, suco, and aldeia. The GCD database was the only dataset to disaggregate location information to the aldeia level, so it was not used for matching purposes. The frequency of displacements made location information difficult for witnesses to pinpoint exactly, except in places where the violation occurred in the place where the witness currently resided or from where they originally were displaced. People may have been displaced multiple times, across multiple locations and because the conflict was spread over three decades, recall of exact locations were subject to a number of errors.

Additionally, the boundaries between geographic locations is affected by three factors – the difference between Indonesian and Timorese place names and geographic divisions over time, the imprecision of boundaries, especially in rural areas, and potential data collection, coding, and entry errors. As a result, matches anywhere within a single district and between bordering districts was considered. Potential matches between a sub-district and suco that were closer to each other were given a higher preference as well. In studying the data closely, records that matched on a preponderance of data fields other than place provided substantiation for our judgments on location matching. Where the HRVD documented a death occurring in the same location as the interview location, it was assumed that the location information was likely to be accurate.

In rare cases, matches were made that violated the rule for location data, but only when it was clear that the records identified the same person, and that common typographical errors accounted for the difference. When there was more than one possible match, the matching algorithm tried to match to the less-specific records in order to preserve more-specific records for later match candidates. When there was equal distribution between locations at any geographic level, the less specific location was preferred and if there one was not more less specific, than one was randomly selected to be the "rep rec".[101]

5.5.1.4 Matching Dates

As the conflict in Timor-Leste was long, many respondents did not remember the exact dates and places in which events occurred. The GCD data was assumed to be more accurate for date and place information because bodies would normally be buried shortly after death, and close to the

[101] The rep rec is the record that best represents that grouping of matched records by having the most complete data. Records with the most common date or place within that group or a record with a more precise place or date is considered more complete. The more complete the data, the better each subsequent round of matching for both intra- and inter-system matching will be. Because records were being linked together and the data unique to each record preserved, as opposed to deleting duplicates, it was necessary to look at the variation within the matched records to see if the differences would significantly change the analysis.

place of death. When matching on the date field, the record linkage expert would link records that were plus or minus 3 years from each other. The exceptions to this rule were rare, and only made when the other data fields were strong exact matches. Records with month and day data were often inaccurate in the HRVD and RMS data as memory tends to be faulty over such a long period. Therefore, more-specific dates were matched to each other where they were close, and to less-specific dates where they were not close.

5.5.1.5 Record Level Constraints

Matching constraints were implemented to prevent over-matching. Specifically, the following matches were not allowed:

i. records of victims from the same statement (because each statement identified unique victims which may have had the same names because of familial relationships

ii. two non-fatalities could not be matched if they were reported in the same source record (because the data coding and database representation methods used prevented duplicate records from a single statement being entered into the database)

iii. a deponent could not match to a fatal violation

iv. a non-fatal record could not match to a fatal record if any dates associated with the non-fatal violations were before the fatal records DOB

v. a non-fatal record could not match to a fatal record if any dates associated with the non-fatal violations were after the fatal records DOD

5.5.2 Intra-System Matching

Within a dataset, a person may be identified by multiple witnesses. Intra-system matching links records that identify the same person to generate a list of unique named persons to prevent over-counting, and thus, over-estimations. Intra-system matching is very complex and difficult to perform in a database as a person can match to n other records in the dataset. Therefore, the data is manipulated in a spreadsheet which makes it easier to order and reorder the data in multiple ways to locate linkages that need to be made.

Intra-system matching a dataset before merging its records with other datasets can reveal patterns inherent in that data collection project. Some of these patterns may be systematic errors in data collection, coding or data entry, or may be the result of the structure of the data collection. The observation of patterns within each dataset allows for the investigation, and if necessary, the correction of the underlying errors.

The three datasets of the CAVR would have been too large to do high quality data matching if combined because some of the patterns would have not have been noticeable to the human eye. That is, if all three datasets were combined into a single list, the resulting list would include more than 160,000 records. Finding matching records in a list this long would have been very difficult for a human reader.

5.5.2.1 HRVD Intra-system Fatal Matching

First, intra-system matching on fatal data in the HRVD was performed to link records that described the same victim. The records were imported into a spreadsheet and sorted on first name, last name, POD, and DOD, to find records that matched.

As records were linked, a “rep rec” was chosen. After each sort, a matching pass was performed and the linked records within a match group hidden (but not dropped) from the outputted data file, leaving just its “rep rec”. This reduced the noise within the data. Noise can be defined as the non- “rep rec” records in a match group that distract the matcher from the potential relationships of the “rep rec” to other candidate matches. The smaller the list of unique records, the easier it is to see potential matches and other patterns within the data. Each subsequent pass identifies additional matches, and finally, a list of unique records is distilled from the entire dataset. A minimum of five passes are done on each dataset.

The 15,043 fatal records of the HRVD dataset were reduced to a list of 11,145 unique victims. All the records are then imported back into the Analzyer data matching system. The matched records were linked back to the ´rep rec¡ for analysis when all matching was completed.

5.5.2.2 RMS Intra-system Fatal Matching

The RMS intra-system fatal matching was performed in a spreadsheet after the HRVD intra-matching was completed. The RMS intra-system matching used the same fields as the HRVD intra-system matching and also looked at the source of the record. Records of fatalities collected from the same household were not allowed to match to each other as they identified unique individuals, even if they shared the same name and DOD.

The 4,883 fatal records of the RMS dataset were reduced to a list of 4,619 unique victims.

The resulting linkages of both the HRVD and RMS datasets were imported back into the Analyzer data model for use in computer-assisted and computer-generated matching, and to generate data for analysis. Information and patterns documented by the record linkage expert in the hand-matching phase was then used to generate matching rules and algorithms for the computer-assisted and computer-generated matching processes.

5.5.2.3 HRVD Intra-system Non-Fatal Matching

Computer algorithms were devised to clean and match non-fatal violations in the HRVD. This step is referred to as auto-matching. Automated matching algorithms for the non-fatal violations in HRVD were developed as time and resource limitations did not permit the use of a human record linkage expert. There were three times as many non-fatal victims as fatal victims reported in the HRVD.

The HRVD contained 41,546 records. The intra-system auto-matching yielded a list of 37,651 unique victims of fatal and non-fatal violations.

5.5.2.3.1 Auto-Canonicalization of Non-fatal Name Values & Matching

The first step in the auto-canonicalization process was to build a table with the different cleaned versions of all (fatal and non-fatal) original names in the database. For the first name, the versions were normalized, normalized-terse, first word of normalized (called first-namefirst), and first word of normalized-terse (called first-namefirst-terse). The same method was applied to the last name, except the last word was used instead of the first word. Then, for each victim name of a non-fatal violation, an attempt was made to match the following combinations of the normalized non-fatal full names to all of the normalized hand-canonicalized full fatal names:

namefirst + namelast

namefirst-terse + namelast-terse

first-namefirst + last-namelast

first-namefirst-terse + last-namelast-terse

The matching program matched on a full set of information before trying to match on less information. This matching of non-fatal to fatal-names was only done for normalized fatal names that mapped to a unique canonical name; as the information became more terse, there were fewer and fewer "allowable" normalized names to match on (which was offset by the fact that it was easier to make the match, because the less-terse information was more resistant to coding variability and data entry errors).

For those full names that could not be canonicalized, the first names and last names were canonicalized independently. The order of matching first names was as follows:

namefirst

namefirst-terse

first-namefirst

first-namefirst-terse

A subsequent matching process was developed to follow the preliminary matching round based on the auto-cleaning and auto-matching processes. This process targeted potential matches with the non-normalized names and identified the information-density per data-field of each name record. The percentage of records that contained non-blank values for the respective data fields was as follows:

9% had date_birth (all of these have birth_geo1)

44% had birth_suco_location

50% had birth_subdistrict_location

53% had birth_district_location

70% had Firstname

94% had Sex

100% had Lastname (since it's a mandatory field required for matching)

Since the last name field was the only non-blank field for all records, it was the only field that could be used in the index blocking. Blocking looks at records where the field(s) being blocked share the same value. The blocking for the last name field was done on the first four letters of each name. The match algorithm had to be carefully calibrated: if there were many blank fields, then a closer match on the non-blank fields was required (also, matches on very common last names were given less weight).

There were three different kinds of “closeness” that were varied:

1. the number of letters in the name that matched (4, 8, or all),
2. the number of levels in the birth location that matched (from 1 through 3), and
3. the required-closeness of the dates (from 1/3 year to 3 years).

With two-thirds of the victim names auto-canonicalized, and a well-defined set of rules for required-closeness-of-match for different numbers of non-blank fields, the resulting match rate was approximately 15% (compared to about 25% for the human matched fatal-violations data).

A match rate of 15% for non-fatal violations seems plausible as:

i. only two-thirds of the name records could be canonicalized, and

ii. we expect higher reporting density for fatal violations – as they are more easily identifiable and easier to recall by a larger number of people in the victim's social network.

The automated inter-system matching on the non-fatals reduced the dataset from 44,203 records to a list of 31,568 unique victim records.

5.5.2.3.2 Data Linkage Expert Review of HRVD Non-fatal Intra-system Matches

The record linkage expert studied a sample of the auto-matched results to make sure there were no obvious mis-matches (i.e. over-matching). No systematic pattern of over-matching was found in the review of a random sample if 10% of the matched group record. The largest group of records which were matched to each other was 20 records. A review was done of the largest groups to ensure that their match size was plausible.

Intra-system matching on fatal data generates a combined list of unique individuals who are all dead, even though the cause of death can vary. When intra-matching is done on non-fatal violations, a victim can suffer one or more violations, on one or more days, in one or more places. The non-fatal matching reveals the human rights violations suffered by individual victims, where a victim may have suffered other violations that may or may not have resulted in a fatality.

5.5.3 Inter-System Matching

Inter-system matching links lists of unique individuals from multiple datasets and is done cumulatively in pairs or datasets. Inter-system matching is applied only to fatal data. First, inter-system matching is applied using the 11,126 intra-system matched records from HRVD to the 4,619 RMS intra-system matched records in the Analyzer Record Linkage application. The RMS fatal source dataset was matched into the HRVD fatal target dataset.[102]

5.5.3.1 Phase 1 - Computer Generated Matching

Strict matching (referred to as P1 matching) automatically identified “exact matches”. Processing of “exact matches”[103] via the automated P1 process eliminates the inefficiency of having a human compare every record in, or between the databases, with every other record.

Matching based on algorithms was applied to the data to generate a list of potential matches that were deemed to be highly probable. Calculations based on probabilities and frequencies of each data field within a record were weighted and ordered by rank, and a threshold level was

[102] The designation of source and target is determined by the number of records in the dataset. The smaller of the two datasets in the pair is the source and the larger is the target. This is to reduce the number or records that have to be compared, but each record from both datasets are compared to all of its potential matches.

[103] An "exact match" is where two or more records in a database are matched together when all the fields on which matching decisions are being made are identical.

established where the match being made was likely correct. The threshold was set after a review was made of the prospective algorithm-based matches, which eliminated the need for a human to compare every record for possible matches. Potential matches below that threshold were handled one of two ways, depending on whether or not matching was for fatal or non-fatal, and intra- or inter-system matching.

For inter-system matching of fatal violations data, the algorithm-generated match pools were imported into the Analyzer data matching system and the record linkage expert reviewed these computer-assisted match targets for each of the remaining unmatched source records. Non-fatal intra-system matching was completely automated with results reviewed by the record linkage expert to ensure that extreme over- or under-matching was not occurring.

5.5.3.2 Phase 2 - Computer-Assisted Matching

Computer-assisted matching, referred to as P2, was based on algorithms that generated pockets of potential matches between source and target records that were deemed to be likely matches, but required human review to select which of the closely weighted records was the best match. Calculations based on probabilities and frequencies of each data field between pairs of records were weighted and ordered by rank based on names, date of birth, date of death, place of birth, and place of death. Using the Analyzer matching interface, the record linkage expert selected which target record from that pocket, if any, matched the source record being examined.

The P2 fatal inter-system matching rules were:

1. The sex of source and target(s) had to be equal, where sex was known.
2. The first initials of names between a source and target(s) had to be the same.
3. For target(s), where DOB and DOD were known, one of the dates had to be within 5 years of the source dates.
4. If the source and potential targets(s) had 'perfect' DOB or DOD, at least one of the other matching fields hand to match.

After the inter-system match work was done in Analyzer between the HRVD and RMS datasets, the resulting list of unique fatal victims was imported into a spreadsheet. The records were then sorted on the various data fields to determine if any other possible matches could be found. This not only served to catch matches missed, it also measured how good the matching algorithms had been. Additional fine tuning of algorithms was done as a result of the hand reviews by the record linkage expert, ensuring that successive matching passes would be more thorough and accurate.

5.5.3.3 Phase 3 – Vague Data Matching

In Phase 3 (P3) matching, records that contained too many blank fields, or were records of commonly-named individuals, from the same area, or who died in the same time period were matched. These matches did not have enough data to be specific about which source/target pair was exact, so one was randomly selected from the targets. For example, Mau Bere was a very common name in many places of the country, and 1999 was a year when many of them died. It is unlikely that there were missed intra-system matches for two reasons. First, they were records that often came from the same testimony statement which indicated they were family members with the same name. Second, the GCD recorded many deaths in the same cemetery with the

same name and date (or no date), but there was not enough identifying information within the HRVD and RMS datasets to distinguish them as unique individuals.

The P3 matching process made matches where equal probabilities of a good match for a record existed, which did not require the judgment of the record linkage expert.

5.5.3.4 Pair-wise Inter-system Fatal Matching

The inter-system matching pair of HRVD and RMS resulted in the new list of unique victims, named the HRVD/RMS dataset. This dataset included 10,594 records found only in the HRVD dataset, 4,087 found only in the RMS dataset, and 532 were found in both HRVD and RMS. These 152,13 total unique records were then inter-system matched with the 149,267 records of the GCD dataset. The HRVD/RMS dataset being the source data and the GCD the target data. The pair-wise matching between the HRVD/RMS dataset into the GCD resulted in 157,000 named deceased persons. This total includes records that were out of mandate or did not have dates of death to verify they died within the mandate period. Only records having dates of death within the mandate period were used for analysis.

The linkages within and between these datasets is used in estimating the total number of dead due to the conflict. Records in this final list can linked back to a single dataset, or a combination of the three datasets. Table 34, below, contains a simple matrix showing the results of the final fatal inter-system matching linkages between the datasets.[104]

Table 34: Matrix showing results of the final fatal inter-system matching linkages between the datasets

	HRVD only	RMS only	GCD only	HRVD & RMS	HRVD & GCD	RMS & GCD	HRVD/ RMS/ GCD	Total	Total
Count	5,203	2,148	141,787	382	5,391	1,939	15	0	157,000
Percent	3.31	1.37	90.31	0.2	4	3.43	1.24	0.1	10

If the intra-system matching caught all possible matches, then only zero or one potential match would have been possible during inter-system matching. Matches may be missed if the records being examined had missing data fields that made it unclear if the two records should have been linked. Human error is also possible when looking at the large quantity of data that was involved in the CAVR project. Generally, we assume a match when a majority of the data fields match, or the records' match weight is within tolerances. If there are not enough fields with complete data, then it is difficult to determine with reasonable certainty whether a record should be included or excluded from matching to another. The latter case was especially true for the very common animist names, like Mau Bere where many people, from the same place, died or were killed at the same time.

After completing the inter-system matching in Analyzer, the data was imported into a spreadsheet for review by the record linkage expert. By looking at the data sorted on different

[104] These are unweighted totals, and they include records with missing dates, out of range dates, missing places, and places outside of East Timor. Out of range records were subsequently eliminated from the analysis.

variables, with multiple processes – both human and automated – we can be confident that all possible matches that should have been made were processed. Additionally, the inter-system matching process may be considered a measure of inter-rater reliability (IRR) because it finds instances where matches were missed in the intra-system phase. By returning to the intra-system data and applying the missed matches, it was possible to not only measure the IRR but also correct the data, producing more reliable data upon which estimates could be based.

Table 35: Inter-System Match Record Count Totals & Percentages for Fatal Violation by Dataset Pair

Step	HRVD to RMS	HRVD/RMS to GCD
Starting Count	HRVD + RMS=HRVD/RMS	
Spreadsheet Matching	Count & Percent	
Adjusted from Missed	Count & Percent	
HRVD/RMS total	Count & Percent	
Starting Count		HRVD/RMS + GCD = MSE
P1 Matching		Count & Percent
P2 Matching		Count & Percent
P3 Matching		Count & Percent
Total Count for MSE		Count & Percent

5.5.4 Reported Pattern of Acts of torture, ill-treatment, threat and property violations over Time

The following graphs, Figures 87, 88, 89 and 90, display the pattern of reported acts of torture, ill-treatment, threat and property violations document in the CAVR's statement-taking process. As with the pattern of reported acts of detention in the CAVR's statement-taking process, reported acts of torture, ill-treatment, threat and property violations do not exhibit any systematic reporting bias.

Figure 87

Source: Database of Narrative Statements Given to the CAVR

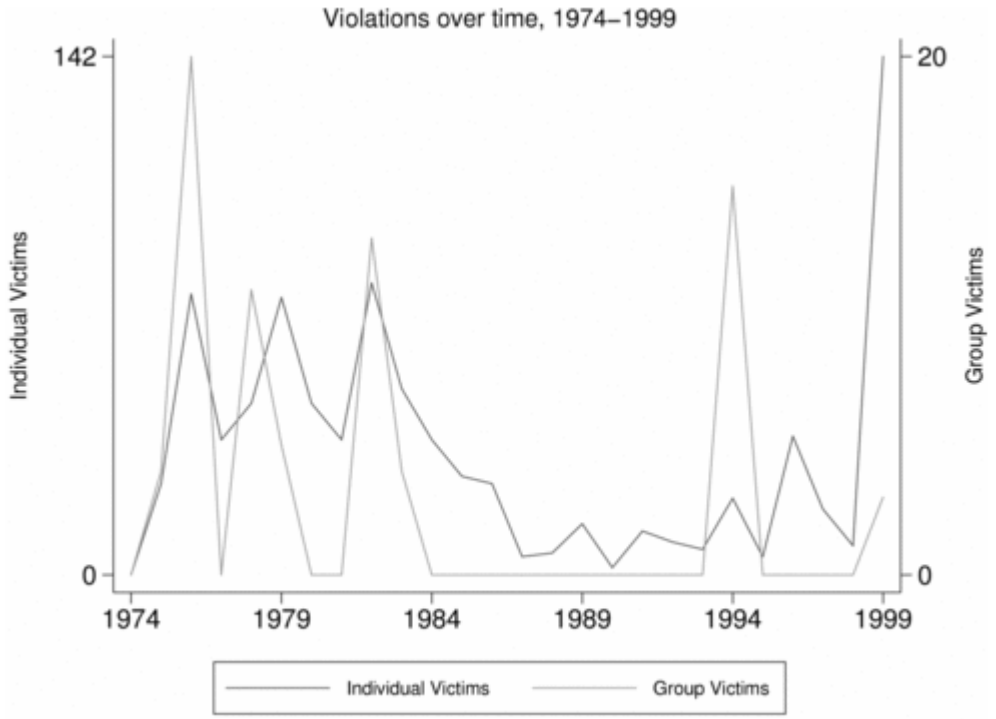
Figure 88

Source: Database of Narrative Statements Given to the CAVR

Figure 89

Source: Database of Narrative Statements Given to the CAVR

Figure 90



Source: Database of Narrative Statements Given to the CAVR

5.6 Data Processing of Reported Violations Involving Groups of Anonymous Victims

During the statement-taking process, a deponent may have talked about one or many victims. Sometimes when multiple victims were reported by a deponent, the deponent did not know some or all of the victims names. In the CAVR statement taking process, 1.9% (1,419/75,443) of victim-records which were documented by the Commission were reported as victims whose individual names were not known by the deponent, but who suffered abuse while they were part of a larger group of people.

In order to integrate this data into CAVR's analysis, and thereby consider violations against named individuals as well as un-named groups, some further processing of the data was required to account for likely duplicate records of violations against a reported victim group. The processing steps to control for this duplication

i. identified violation records (against un-named group victims) which appeared to describe the same victim group, and then

ii. chose a victim record from the pool of possible duplicate records to be retained as the rep rec of this reported victim-violation

Unlike data on violations against individuals (which by-and-large contain personal identifiers such as names, ages, sex etc), violations reported against groups do not usually contain detailed identifiers of the victim-group. As a result, group-victim records were matched together by comparing the following variables of each reported violation against a group:

i. the district where the violation reportedly took place,

ii. the violation-type into which the violation was coded, and

iii. the year and month in which the violation reportedly occurred.

Then after all the like group-victim records were matched together to form a cluster, the record with the largest group-size within each cluster was retained. All other records were regarded as duplicate records and therefore dropped from the dataset.

The level of duplication amongst group-victim records is shown in Table 36. This tables shows how many duplicate violation copies per violation type were identified in the dataset and the number of surplus group violation records which were dropped for the Commission's analysis on violations against group victims.

Table 36: Detention Torture Ill-Treatment Displacement Other Violations All Violations

	Detention		Torture		Ill-Treatment		Displace-ment		Other Violations		All Violations	
Copies	**Obs**	**Surplus**	**Obs**	**Surplus**	**Obs**	**Surplus**	**Obs**	**Surplus**	**Obs**	**Surplus**	**Obs**	**Surplus**
1	441	0	134	0	121	0	180	0	736	0	1,612	0
2	150	75	26	13	30	15	68	34	206	103	480	240
3	69	46	15	10	9	6	21	14	87	58	201	134
4	56	42	4	3	8	6	16	12	60	45	144	108
5	25	20	0	0	5	4	10	8	30	24	70	56
6	6	5	0	0	6	5	12	10	12	10	36	30
7	0	0	0	0	7	6	0	0	0	0	7	6
8	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0
12	12	11	0	0	0	0	0	0	12	11	24	22
13	13	12	0	0	0	0	0	0	13	12	26	24
Total	**772**	**211**	**179**	**26**	**186**	**42**	**307**	**78**	**1,156**	**263**	**2,600**	**620**

5.7 Statistical Estimation Techniques used in the Analysis of Fatal Violations and Displacements

This section presents the survey-based estimation techniques and multiple systems estimation methods used to make the estimates of the total extent and pattern of mortality and displacement during the Commission's reference period.

5.7.1 RMS weight calculations

The survey sampling was described earlier: in 2003, the CAVR field teams interviewed 1396 households selected from 138 aldeieas and groups of aldeieas, called clusters. The clusters were selected by a method called “Probability Proportional to Size” (PPS), and then 10 (or 20) households were selected by simple random sampling in each cluster. If each cluster had exactly the same number of sampled households, the sampling probability of each household would be

identical, a process known as "self-weighting."[105] Due to sampling 20 households in multi-aldeia clusters and non-response in other clusters, not all clusters had the same number of sampled households; however, 78.5% of the sampled clusters have exactly 10 sampled households. Non-response was 3.1%, and so no non-response adjustment was made. The weights were calculated as follows.

For each cluster, the adjustment for varying cluster size is:

$$\text{cluster adjustment} = \frac{\text{median cluster size}}{\text{cluster size}}$$

The raw 1990 household sampling probability is:

$$sp_{1990} = \frac{\text{total number of sampled HHs}}{\text{total HHs in 1990}} = \frac{1{,}396}{168{,}858}$$

and so, for each cluster, the pps weight is:

$$\text{pps\_wt\_1990\_raw} = \left(\frac{1}{sp_{1990}}\right) \cdot \text{cluster adjustment}$$

There was considerable population change due to migration and growth between 1990 and 2004 when the survey was conducted. Before the weights could be estimated, the total number of households in each aldeiea was adjusted from the 1990 census using data from the 2004 census. During the sample design, the clusters were chosen using the household counts for each aldeiea reported by the 1990 census. At the time these calculations were done (April 2005), the Census Timor-Leste 2004 enumeration data were available disaggregated only to the subdistrict level, but not by suco or aldeiea.[106] Note that the 1990-2004 weight adjustments do not affect the total summed weight, which is fixed at the number of households that existed in 2004. The weight adjustments affect how much households in different places affect the projection.

Two subdistricts from 1990 were not listed in the 2004 census results: Fatu Maca in Baucau was absorbed by Baucau subdistrict, and in Oecussi, Pante Macassare B was subsumed in Pante Macassare. For these subdistricts, the number of households in 2004 was estimated by using the proportion of households in the absorbing and absorbed subdistricts in 1990 multiplied by the total in the absorbing subdistrict in 2004.

Although the 2004 household totals are available from the census at the subdistrict level, the RMS has too few responses at the subdistrict level for the estimates of weights by subdistrict to have adequate data (29 of the 59 sampled subdistricts have fewer than 20 responses). Therefore the 1990 weights were scaled to the 2004 district totals by the following calculation:

$$\text{district adjustment} = \left(\frac{\text{total HHs in 2004 in this district}}{\text{Total 1990 weight in this district}}\right) pps\_wt\_2004 = pps\_wt\_1990\_raw \times \text{district adjustment}$$

By forcing the weights to match the 2004 census's district household counts, the weights were normalized to sum to the total number of households in 2004 (194,943). The errors given in the results are calculated using Stata's standard survey modules.[107] These modules use the survey design variables (stratum, primary sampling units, and sampling weight) to make weighted

[105] See Paul S. Levy and Stanley Lemeshow, Sampling of Populations, New York: Wiley, 1999, ch. 11.
[106] See http://dne.mopf.gov.tp for the census data.
[107] Stata Corporation, Stata Survey Data Reference Manual, v. 8: College Station, TX: Stata. 2003.

estimates of the totals and Taylor-series approximations of the sampling errors. The error estimates assume random sampling with unequal sample weights. This assumption is conservative (i.e., it will tend to overestimate the sampling error) with respect to weights calculated using the PPS methods described above.[108] The data files used for these calculations are available at http://www.hrdag.org/timor.

5.7.2 RMS date assignment for displacement analysis

The survey asked respondents when they moved from each of their locations during the period 1974-1999. When respondents were uncertain of the specific date of their move, they often identified the year of the move with a period of the agricultural cycle or whether it was the dry or rainy season. For each of these partial or seasonal dates, we assigned the displacement to the quarter in which the period or season fell. Where the partial date identification could fall in more than one quarter, it was randomly assigned to a quarter. Of the 2024 moves defined by the respondents as displacement events, 76.6% were identified at least to the quarter, and 15.7% more were identified by the season. Only 7.7% of the displacement events were identified by year without specifying the month.

5.7.3 RMS weight adjustments for mortality estimates

The calculation of the weights assumes that events reported by each household could only have been reported by that household. This assumption is the result of the weights being simply the reciprocal of the sampling probability for the given household. Therefore, if there were more than one household that could have given information about a specific death, the true sampling probability for that death is greater than the probability for a single household. Deaths reported by the survey respondents violate the single-reporting-household assumption because for each death, there may have been more than one household which could have given information about that death. Among the 5,402 total deaths reported by respondents, 545 were reported more than once (the duplicate reports were identified and removed before estimation). The duplicate reporting implicit in the survey weighting was corrected by adjusting the weights as described below.

Before the survey weights can be used to estimate the total number of deaths, they must be adjusted to account for the number of households that were potential respondents for each death. That is, for each death, how many relatives survived until 2003 to be potential respondents in the survey? Much of the information required for this calculation is available in the survey because the respondent's relatives are also the decedent's relatives. The number of surviving relatives for each decedent D was calculated based on the relatives reported by the respondent R using the following rules:

1. If D is a parent of R, the expected number of relatives surviving in 2003 is the sum of the following:
    a. assume that D's parents are 25 years older than D (or 50 years older than R, if D's age is not reported); use age-specific conditional probabilities of survival

[108] See Donna Brogan, "Sampling error estimation for survey data" Chapter XXI in Household Sample Surveys in Developing and Transition Countries, United Nations Publication ST/ESA/STAT/SER.F/96, Department of Economic and Social Affairs of the United Nations Secretariat, 2005.

(calculated from the survey) to estimate the expected number of parents alive in 2003;

b. count R's siblings as D's children;

c. given an average approximate total fertility rate of 5 prior to 1975, assume that D had 4 siblings with ages (-4, -2, +2, +4) years from D's age (if D's age missing, set D's age to R's age + 25), calculate the siblings' ages in 2003, and multiply each by the conditional probability of surviving to that age, and sum

2. If D is a sibling of R,

   a. D's parents are R's parents, count the survivors directly;

   b. R's siblings are D's siblings, count the survivors directly;

   c. assume that D had the same number of surviving adult children as R.

3. If D is a child of R,

   a. R and spouse are parents, count the survivors directly;

   b. adult children of R are D's siblings, count the survivors directly;

   c. assume no surviving adult children of D.

This calculation yields the expected surviving adult relatives for each D, as well as indicating which of these surviving relatives live in R's household, and which live in other households.

To convert the expected surviving adult relatives of D into an adjustment for the sampling weight, the number of relatives must be converted to an expected number of households in which the relatives live. There are on average 0.5 relatives of D (in addition to R) living in R's household. Assume that other households in which D's relatives live have the same concentration of relatives per household as R's household (i.e., 1.5 relatives per household). Thus, if D has L surviving relatives who live outside of R's household, there are $a = 1 + \frac{L}{1.5}$ households which could give information about D. The survey weights were adjusted for possible multiple reporting of D by dividing each D's sampling weight by this factor, *a*. This calculation assumes that the other potential respondent households are in the same cluster as R, or that they are in a cluster with a similar within-cluster sampling probability.

5.7.4 Sensitivity analysis of assumptions in mortality reweighting

There are a number of assumptions in the weight adjustments for the mortality estimates, including the following:

- The period difference between generations (assumed to be 25 years)
- The number of siblings respondents' parents had (assumed to be 4)
- The birth spacing of parent's siblings (assumed to be 2 years)
- The number of adult children respondent's siblings had (assumed to be equal to the respondent's children)

These assumptions were tested using the following variations, and the annual total number of deaths were calculated:

- The inter-generational spacing was varied to 18 and 30 years
- The number of siblings respondents' parents were assumed to have was increased to 6
- The birth spacing was increased to 5 years between siblings
- The number of adult children respondent's siblings had was assumed to be double the number of the respondent's children

For each variant estimation, the annual totals were tested (by a two-mean t-test) against the main model. None of the years in any of the variant models was significantly different at $p<0.05$. The minimum p-value was 0.13, and it was an outlier: the second-lowest p-value was 0.23. Therefore, the estimates are not substantially sensitive to the assumptions about family structure.

Although the estimates are robust to the assumptions about family structure used to estimate the number of surviving relatives who could give information about D, the magnitudes of the estimates are sensitive to the model used to transform the estimated surviving relatives to estimated households that contain relatives. The estimated number of surviving relatives is L, and the estimated number of households containing relatives of a decedent D, denoted *a*, is $a = 1 + \frac{L}{1.5}$. The denominator 1.5 comes from the average number of relatives for D (including R) living in R's household (0.5). Varying this average from 0 to 3 (i.e., assuming 1 - 4 surviving adult relatives per household) varies the resulting estimates of the total estimated deaths (by all causes) from -14.2% to +19.6%. The effect of varying this model declines over time, with the largest variations found in the early years 1972-1975 (-21%, +26%) and the smallest variations found in more recent years 2001-2003 (-11%, +16.2%). The decline is consistent over time.

Given a constant number of surviving relatives, fewer surviving relatives per household means more potential reporting households, a higher estimated sampling probability per reported death, and a lower sampling weight per reported death, and therefore fewer estimated total deaths; more adults per household reverses this logic.

Although the total estimates vary with changes in the model transforming relatives into households, the patterns are constant. The correlation coefficients for the main model to the low (0) and high (3) models above are each 0.99. Although the model of relatives-per-household does affect the total magnitude of the estimated deaths, it does not affect the estimated patterns over time.

5.7.5 Multiple systems estimation (MSE) motivation and theory

The survey analysis is conservative in the sense that it corrects for potential duplicate reporting by matching deaths across households, and because there is an adjustment to the sampling weights based on the estimated number of households which could have given information about each death. As some deaths may be reported by several households, there are other deaths which occurred during 1974-1999 for which there are no surviving relatives in 2003. If entire households died during the mandate period, there would have been no collineal relatives who could have given information in 2003. Given these limitations, an alternate method for estimating the total deaths may provide a check on the survey estimates.[109]

[109] This explanation follows P. Ball, J .Asher, D. Sulmont, D. Manrique, "How many Peruvians have died? An estimate of the total number of victims killed or disappeared in the armed internal conflict between 1980 and 2000, a

MSE uses several separately-collected incomplete lists of the population. The lists are matched identifying the elements common cross lists in order to estimate the number of elements that are missing from all of the lists. In this project, deaths documented in the HRVD, RMS, and GCD were matched across the three systems using the name, date of death, location of death, and date of birth.

The most basic form of this technique is capture-tag-recapture, which uses only two lists.

A technical explanation of how a count of the unknown members of the population can be estimated is as follows. Consider the case of two projects P1 (a list of A individuals) and P2 (a list of B individuals). There are M individuals who are matched across both lists, in a universe of N total individuals (N is unknown). If all of the people in the universe N have an equal probability of appearing in List 1, then the probability of a specific individual being reported by P1 is

$$\Pr(\text{captured in list 1}) = \frac{A}{N}$$

Similarly, if all of the people in universe N have an equal probability of appearing in List 2, then the probability of a specific individual being reported by P2is

$$\Pr(\text{captured in list 2}) = \frac{B}{N}$$

The probability of a specific individual being captured in both lists is

$$\Pr(\text{captured in list 1 and list 2}) = \frac{M}{N}$$

By definition, the probability of an event composed of two independent events is the product of the independent probabilities. Therefore,

$$\Pr(\text{captured in lists 1 and 2}) = \Pr(\text{captured in list 1}) \times \Pr(\text{captured in list 2})$$

Which is $\frac{M}{N} = \frac{A}{N} \times \frac{B}{N}$: given this equation, solve for *N*. Rearranging the terms, $\frac{M}{N} = \frac{A \times B}{N \times N}$ and then multiplying by N, $M = \frac{A \times B}{N}$, multiplying again $M \times N = A \times B$, and finally dividing by *M* yields $N = \frac{A \times B}{M}$. Note that with the final equation, the total number of deaths *N* can be estimated using the totals from *A* and *B* and from the matches between them, *M*.

There are many assumptions implicit in this solution. For example, none of the lists have individuals reported twice and that matching between the lists is accurate. In this project these two assumptions were controlled during the data processing as described in the matching section.

Other assumptions inherent in the capture-tag-recapture model are more difficult to manage. First, the method assumes that individuals are not entering or leaving the universe during the process of creating the lists, and second that the lists were selected randomly from the population. In human rights documentation projects, the first assumption is usually irrelevant because the documentation occurs retrospectively. The second assumption cannot be satisfied, and it must be replaced by the assumption that the estimation is robust to the selection process.

---

report to the Peruvian Truth and Reconciliation Commission." Washington, DC: AAAS. 28 August 2004. Available online at http://shr.aaas.org/hrdag/peru.

Another assumption is that the lists are independent, that is that the probability that an individual is in list two is independent of the probability that the individual is captured in list one. The final assumption is homogeneity: that the individuals that compose the universe all have the same probability of being captured.

If either of these assumptions is violated, the capture-tag-recapture method will not yield an adequate estimate of the total population size. If there are more than two lists with adequate information, the problems of dependency or heterogeneity can often be managed through the specification and selection of appropriate models. However, in the data for the HRVD, RMS, and GCD, there are only two usable systems (RMS-GCD for deaths due to hunger and illness, and HRVD-GCD for killings).[110] Alone, these estimates would be insufficient, but in combination with the RMS estimates, these estimates provide useful additional information.

5.7.6 Allocating GCD by type of death

The graveyard data does not include the manner of death. There were 89,894 graves with at least a first initial (or name), a last name, and a year of death between 1972-2003. Of these, 7,117 matched either the HRVD or the RMS (or both), and through this match, the manner of death can be learned from the matched record's manner of death. The remaining 8,2717 GCD records need to be allocated to the four categories of manner of death (killing, death due to hunger and illness, combatant deaths, and other deaths). From the RMS, annual proportions of deaths by these four types are shown in Table 37, below. Note that these proportions exclude deaths for which the manner of death is unknown (204 of 3,235 deaths reported in the RMS between 1969-2004 have unknown manner of death).

[110] The initial application of multiple-systems estimation to demographic estimation was by C. Chandra Sekar and W. Edwards Deming, "On a Method of Estimating Birth and Death Rates and the Extent of Registration," Journal of the American Statistical Association, March 1949: 101-115. A thorough discussion of the estimators for the dual-system approach and the relevant error calculations is available in Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul H. Holland. Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA: MIT Press. 1975. For commentary on the use of these methods in human rights analysis, see Fritz Scheuren, "History Corner," The American Statistician, February 2004.

Table 37: estimated proportions of deaths, by period and manner of death

Period	Killing	Hunger/ Illness	Combatant	Other
1972-1974	0.9%	95.9%	0.0%	3.2%
Margin of error	1.8%	5.1%	0.0%	4.9%
1975-1982	11.2%	83.0%	4.4%	1.4%
Margin of error	4.7%	5.1%	2.5%	0.6%
1983-1998	5.5%	86.5%	0.7%	7.2%
Margin of error	2.5%	3.7%	0.6%	2.5%
1999	16.2%	83.0%	0.4%	0.4%
Margin of error	10.2%	10.2%	0.8%	0.8%
2000-2003	3.5%	86.9%	0.8%	8.9%
Margin of error	3.1%	6.5%	1.6%	4.9%
Total	**8.3%**	**85.1%**	**2.4%**	**4.3%**
Margin of error	**2.7%**	**3.1%**	**1.2%**	**1.2%**

These proportions were used to allocate the unmatched GCD records to the distinct manners of death to be used in the MSE calculations for each year: the proportions from the period containing each year was used to allocate the GCD deaths in that year. The margin of error of the allocation was included in the estimated error for the MSE estimates.

5.7.7 Sensitivity analysis of the loss of social knowledge: adjustments for underestimates

The survey asked respondents about the deaths of their parents, siblings, and children. However, some deaths left no parents, siblings, or children who survived until the survey was conducted in 2004. If deaths occurred long in the past, even the decedents' children would have all died, leaving no one to report the deaths. In other cases, small families may have suffered complete mortality, so that no one survived to report the deaths. As the survey estimates the number (or the rate) of deaths farther back in time, the underestimate resulting from the loss of social knowledge must become more severe. However, even in the nearly immediate past (for example, in 2003 for a survey conducted in 2004), it will be impossible to document some deaths which have left no survivors. For example, people who have no surviving parents, siblings, or children who died in 2003 cannot be reported in the survey.

The crude death rate (per 1,000 people) is an estimate of how many people died, in total, by year. It is a standard demographic and health indicator, usually estimated by indirect methods using census records. For East Timor, these rates are difficult to estimate because the quality of the 1980 and 1990 census data has been in dispute.[111] The CDRs estimated by the US Bureau of the Census for East Timor are shown for 1990-2004. The Indonesian overall rate is shown for 1983. The estimate shown for 1971 comes from an Indonesian government claim that in all of

[111] See, e.g., Kiernan, Ben "The Demography of Genocide in Southeast Asia: The Death Tolls in Cambodia, 1975-79, and East Timor, 1975-80." Critical Asian Studies 35:4 (2003), 585-597.

Indonesia between 1971-1990, the CDR declined by 45%; the 1971 estimate shown here is the 1990 estimate for East Timor inflated by this factor. A projected CDR is also shown by linearly interpolating between the 1971 estimate and the 1990-2004 estimates.

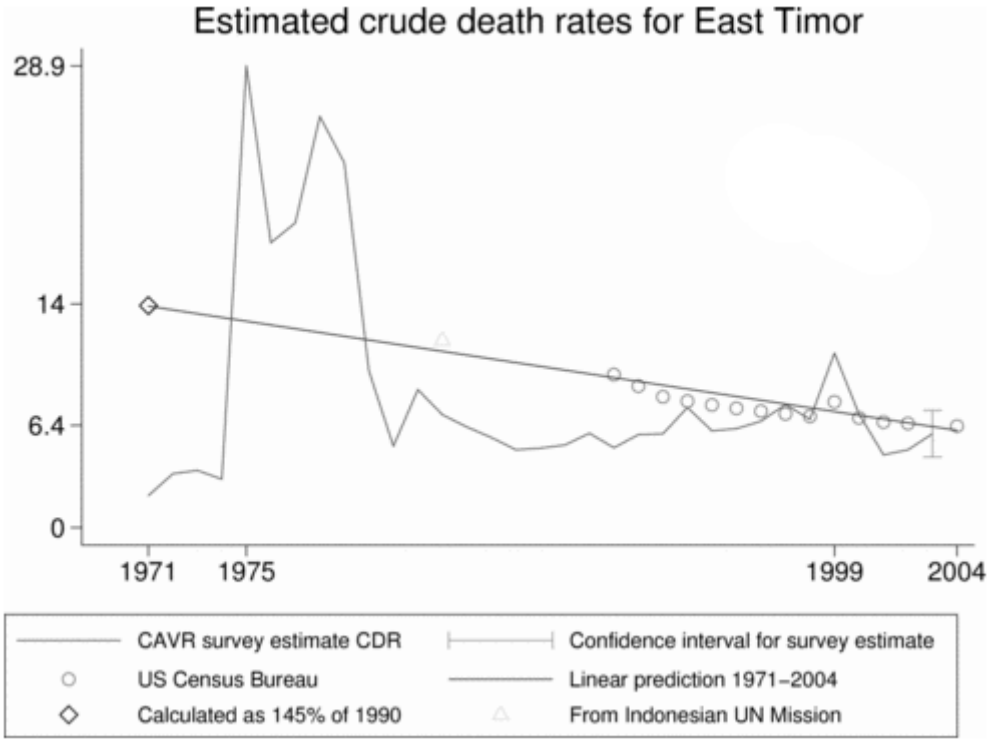
In addition to the CDR estimates, the CDR from the CAVR's RMS is shown. This estimate is the total estimated deaths divided by the estimated population for that year (multiplied by 1,000). There are several useful observations from this graph. First, the CDR estimated by the US Census Bureau is within the confidence interval of the CDR estimated by the RMS beginning in 1993. In 2003, the confidence interval of the RMS CDR (4.2 – 6.6) contains the US Census Bureau estimate (6.4), as shown in the graph by the capped spike at the end of the CAVR line. That is, while the RMS greatly underestimates the death rate in the "normal" peacetime years 1972-1974, by the mid-1990s, the RMS agrees with the results obtained via the indirect methods employed by the US Census Bureau. This observation is consistent with the notion that the RMS estimates suffer increasing downward bias into the past.

During years in which the historical record suggests that substantial excess deaths occurred, the linear interpolation of the CDR underestimates deaths. These years include 1975-1979 and 1999.This is consistent with the literal meaning of "excess" deaths. (There are no census-based CDR estimates for the 1975-1979 period) Looking farther into the past, the survey-based CDR captures a decreasing fraction of the total CDR (a similar graph can be drawn for the MSE estimates over time, with similar results).

To adjust the RMS, the deaths lost to the loss of social knowledge must be estimated over time. The model employed was the following:

- the number of deaths estimated by the CDR and the projected population for each year was estimated (CDR_deaths), shown as a rate in Figure 91;

Figure 91



- the fraction of CDR_deaths that occurred due to hunger and illness was estimated using the fraction of all deaths reported in the survey that were due to hunger and illness (similar to the allocation used for the unmatched GCD data). In the survey, the mean (and median) fraction of all deaths (over years) attributed to hunger and illness is 0.80, and 50% of all years are within the range 0.754 – 0.846;
- the ratio of estimated deaths to CDR_deaths was calculated for the peacetime years (1972-1974 and 2002-2003); this is the fraction of "rememberable deaths," called the "memory fraction;"
- The memory fraction for 1975-2001 was estimated by linear interpolation using the following equations:

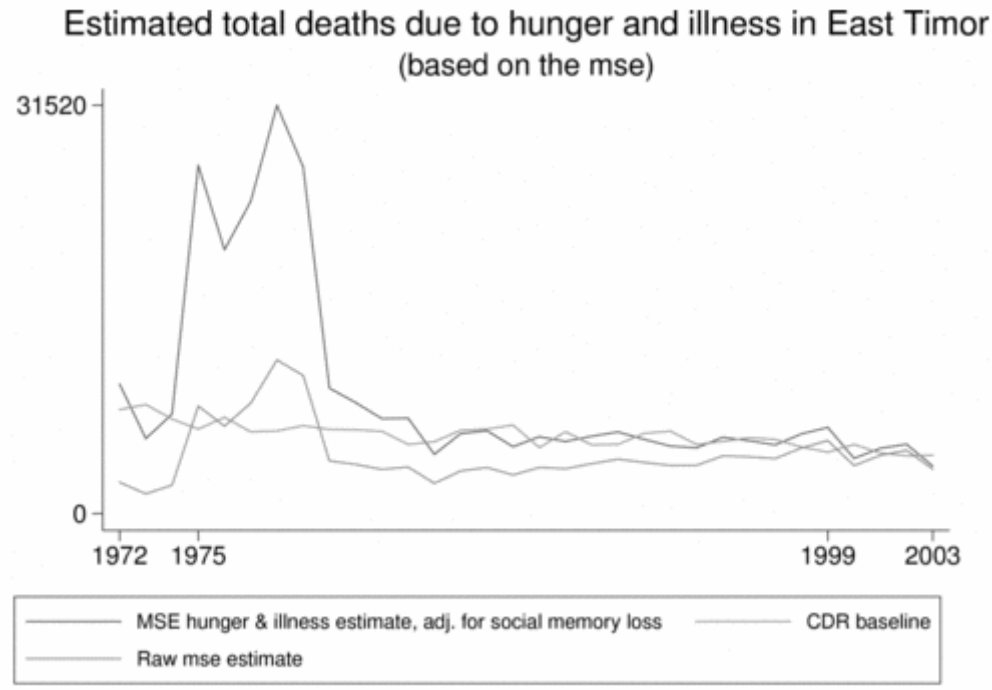
  estimated memory fraction (MSE) = -39.1 + 0.0200*year

  estimated memory fraction (RMS) = -43.9 + 0.0224*year

- The memory fractions for MSE ranges from 0.241-0.936, whereas for the RMS, they ranged from 0.228-0.846. This difference has an enormous impact on the outcome.
- The adjusted estimate was calculated as the original estimate divided by the memory fraction for each year.

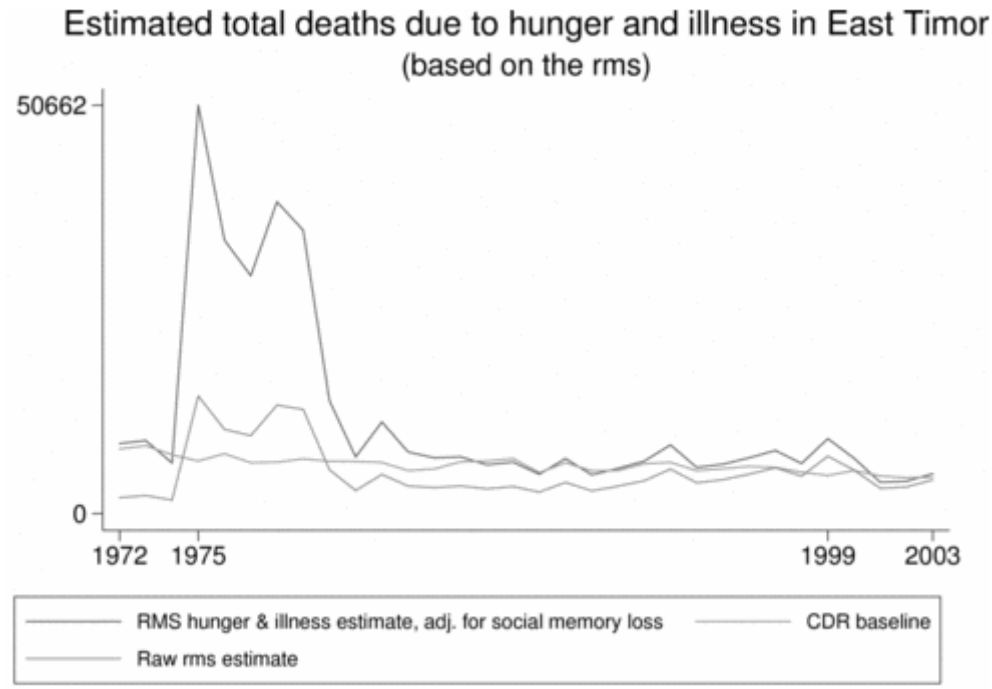
The adjusted estimates are presented below in Figures 92 and 93. Note that in both graphs, the raw estimates and the adjusted estimates converged as the year approached 2003. The impact of the higher memory fraction for the MSE relative to the RMS was apparent in the estimated total deaths in excess of the CDR baseline: the MSE adjusted estimate was 104,000 deaths while the RMS adjusted estimate was 183,300 deaths.

Figure 92



Estimated total deaths in excess of the CDR baseline= 103904

Figure 93



Estimated total deaths in excess of the CDR baseline= 183332

Both of these estimates depend on a number of assumptions, including assumptions about the shape of the decline of the CDR from the early 1970s through the late 1990s and about the nature of the loss of social memory. Smooth but non-linear changes in the loss of social memory (either

concave up or concave down) would not change the estimate substantially. However, if the underestimates in the MSE and RMS due to social memory loss were somehow discontinuous or otherwise drastically different for 1972-1974 relative to the peak years 1975-1979, the adjustment employed here would not correct appropriately for the underestimate. Both of these models depend on CDRs calculated from the 1980 and 1990 census data and indirect methods used by the US Bureau of the Census. There is sampling and non-sampling error which is not represented in the graphs or the statistics, but the error is certainly substantial.

However, these models have the benefit of showing that with the adjustment, the estimated annual total deaths due to hunger and illness closely match the CDR baseline deaths for the pre-invasion period (1972-1974) and for the period 1984-1998.

There are several reasons to prefer the MSE estimate to the RMS estimate. Although the RMS more closely matches the CDR deaths estimate in the post-occupation years that approach peacetime, 2002-2003, the MSE more closely matches the pre-occupation CDR total deaths estimates. For the purposes of this estimate, the most relevant period is 1975-1979, and the choice of estimates should be guided by the best fit immediately prior to this period. A second reason to prefer the MSE is that it is based on considerably more data than the RMS alone: the MSE uses the GCD data in addition to the RMS.

The strongest conclusion which can be made is that the unadjusted RMS and MSE estimates must be too low. It is possible to argue that total deaths due to hunger and illness in excess of a CDR baseline could be as high as 183,300. However, this high-end estimate requires considerable speculation and is subject to substantial error, and the authors do not recommend it. In our opinion, the appropriate and conservative finding is that there were at minimum, 100,000 deaths in excess of the peacetime baseline due to hunger and illness.

5.8 Retrospective Mortality Survey (RMS) Questionnaire

The English Version of the RMS Questionnaire is reproduced in this section[112].

[112] The Bahasa-Indonesia and Tetun versions of the RMS Questionnaire are on-file with the Commission for Reception, Truth and Reconciliation (CAVR) in Dili.

COMMISSION ON TRUTH, RECEPTION AND RECONCILIATION (CAVR)

RETROSPECTIVE MORTALITY SURVEY [TIMOR-LESTE]
DECEMBER 2003
HOUSEHOLD QUESTIONNAIRE

CONFIDENTIAL

SECTION 0: HOUSEHOLD SCHEDULE

Subsection 0.1: Identification

		Name	Code	
0.1.1	District (DIST)			
0.1.2	Subdistrict (SUBD)			
0.1.3	Suco (SUCO)			
0.1.4	Aldeia (ALDA)			
0.1.5	Household Number (HHNO)			
0.1.6	Name of Household Head (NOHH)			

ELEV	
LS	
BT	

Subsection 0.2: Interviewer Visits

		1	2	3
0.2.1	Interview Date (dd/mm/yy) (DATE)	/ /	/ /	/ /
0.2.2	Interviewer's Name (INNM)			
0.2.3	Interviewer's ID Code (INCD)			
0.2.4	Interviewer's Gender (INGN)			
0.2.5	Supervisor's Name (SPNM)			
0.2.6	Result Code* (RSCD)			
	Interviewer's Signature			
	Supervisor's Signature			

Subsection 0.3: Field notes

		Field Supervisor	Field Supervisor ID	QC Officer	Keyed by
0.3.1	Name				
0.3.2	Date				
0.3.3	Comments				
	Signature				

Subsection 0.4: Re-interview

		Interviewer	Interviewer ID	Field Supervisor	Field Supervisor
0.4.1	Name				
0.4.2	Date				
0.4.3	Comments				
	Signature				

*Result Codes

1	Completed	6a	Unable to complete – Interrupted
2	No HH member at home or no competent respondent at home at time of visit	6b	Unable to complete – Overcome by Emotion
3	Entire HH absent for extended period	6c	Unable to complete – Safety concerns
4	Postponed	6d	Unable to complete – Other
5a	Refused – lack time	7	Dwelling vacant or address not a dwelling
5b	Refused – opposed to study	8	Other (specify)
5c	Refused – other		

Interview Start Time:_____

COMMISSION ON TRUTH, RECEPTION AND RECONCILIATION (CAVR)

RETROSPECTIVE MORTALITY SURVEY [TIMOR-LESTE]
DECEMBER 2003
HOUSEHOLD QUESTIONNAIRE

CONFIDENTIAL

SECTION 1: INTRODUCTION

Hello. My name is [_____]. I am working with the Commission on Truth, Reconciliation and Reception (CAVR). CAVR is a national commission, mandated by statute, which has been set-up to: (1) establish an accurate historical record of the period from 1974-1999, (2) facilitate reconciliation amongst and between communities at both a local and national level and (3) provide the government and civil society with recommendations on addressing past injustices during this period. As part of its efforts to uncover the truth behind what exactly has happened in East Timor between 1974 and 1999, CAVR is conducting a survey of a sample of Timorese to assess the mortality patterns across the country for the different time-periods of the Indonesian occupation. We realize that many people have suffered greatly during this time and may have much to tell. However this survey requires brief responses to a number of prepared questions from the adult members of your household. Your cooperation and participation in this survey will greatly assist CAVR in establishing a reliable and accurate account of the total number of fatalities between 1974 and 1999 in East Timor.

The findings of this study will be used to establish an accurate historical record of Timor Leste's recent history. While we will need to ask you for your name and the names of the members of your household and family, these will remain private and confidential. The individual names and details of your responses will NOT appear in CAVR's final report nor will they be shared with any other organizations. Rather your identity and that of your family and household members will remain confidential. However, the information you provide will contribute to CAVR's efforts in establishing the overall situation that Timorese experienced between 1974 and 1999. Although the CAVR wants perpetrators brought to justice, participation in this survey does not guarantee compensation for losses/deaths experienced by the household, nor does it mean that the individuals in the household will be able to testify at trials or bring specific charges against anyone. Can I please ask you and the other adult members of your household some questions?

SECTION 1.2: DOCUMENTATION

Before we begin the survey, we would like for you to gather some documents for us. If you have any marriage certificates, birth certificates, baptismal papers, or United Nations registry cards here in your house, please get them now so we can refer to them as we ask you questions.

SECTION 2: HOUSEHOLD REGISTER PAGE 1

Now I would like some information about the people who usually live in your household:

	Usual Residents in Household			SEX		Status in Family	AGE	ELIGIBILITY		Family Head
								Male	Female	
No.	2.1	2.2	2.3	2.4		2.5	2.6	2.7	2.8	2.9
(NO CO)	*Please give me the names of ALL persons who usually live in your household, starting with the head of household [and including yourself].*			*Is (name) male or female?* (GEND)		Relation-Ship Code*: (RLCD)	I *How old is (name)? (CAGE)*	Circle line number of all males aged 20 or older	Circle line number of all females aged 20 or older	*Who is the head of the family?* (FMHD)
	First name (FNME)	Last name (LNME)	Nickname (NNME)	M = 1	F = 2					
01				1	2		Exact? Yes No _____	01	01	01
02				1	2		Exact? Yes No _____	02	02	02
03				1	2		Exact? Yes No _____	03	03	03
04				1	2		Exact? Yes No _____	04	04	04
05				1	2		Exact? Yes No _____	05	05	05
06				1	2		Exact? Yes No _____	06	06	06
07				1	2		Exact? Yes No _____	07	07	07
08				1	2		Exact? Yes No _____	08	08	08

SECTION 2 CONTINUED ON NEXT PAGE

INTERVIEW FORM # _____________ CAVR - RETROSPECTIVE MORTALITY SURVEY

SECTION 2: HOUSEHOLD REGISTER CONTINUATION SHEET NUMBER (circle): 1 2 3 4 5 6

	Usual Residents in Household			SEX		Status in Family	AGE	ELIGIBILITY		Family Head
								Male	Female	
No.	2.1	2.2	2.3	2.4		2.5	2.6	2.7	2.8	2.9
(NOCO)	*Please give me the names of ALL persons who usually live in your household, starting with the head of household [and including yourself].*			*Is (name) male or female?* (GEND)		Relation-Ship Code*: (RLCD)	*How old is (name)? (CAGE)*	Circle line number of all males aged 20 or older	Circle line number of all females aged 20 or older	*Who is the head of the family?* (FMHD)
	First name (FNME)	Last name (LNME)	Nickname (NNME)	M = 1	F = 2					
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐
☐☐				1	2		Exact? Yes No ______	☐☐	☐☐	☐☐

If Additional Continuation Sheet Used, Check Here: ☐

INTERVIEW FORM # _____ CAVR - RETROSPECTIVE MORTALITY SURVEY
SECTION 2: HOUSEHOLD REGISTER PAGE 2

Now I would like to ask a few more questions about your household.	YES	NO	DK
2.10. *Are there are any other persons in your household such as small children, infants, elderly persons that we have not listed?* IF YES, GO TO PREVIOUS SHEET OR CONTINUATION SHEET AND LIST THESE PEOPLE, THEN GO TO NEXT QUESTION.	☐	☐	☐
2.11. *Are there any other people who may not be members of your family, like servants, friends, lodgers, but who usually live here?* IF YES, GO TO PREVIOUS SHEET OR CONTINUATION SHEET AND LIST THESE PEOPLE, THEN GO TO NEXT QUESTION.	☐	☐	☐
2.12. *Are there any other guests or visitors who have been seasonarily staying with you for the past 6 months?* IF YES, GO TO PREVIOUS SHEET OR CONTINUATION SHEET AND LIST THESE PEOPLE, THEN GO TO NEXT QUESTION.	☐	☐	☐
2.13. *Are there any persons who usually live here who have been away for less than 6 months?* IF YES, GO TO PREVIOUS SHEET OR CONTINUATION SHEET AND LIST THESE PEOPLE, THEN GO TO NEXT QUESTION.	☐	☐	☐
2.14. *Are there any persons who we have listed who have been away for the past 6 months?* IF YES, GO TO PREVIOUS SHEET OR CONTINUATION SHEET AND CROSS OUT THESE PEOPLE.	☐	☐	☐

STOP. USING THE ORIGINAL AND CONTINUATION SHEETS, PICK A RANDOM ADULT MEMBER OF THE HOUSEHOLD TO ANSWER MODULES 4, 5, and 6. PICK A RANDOM FEMALE ADULT MEMBER OF THE HOUSEHOLD TO ANSWER MODULE 7.

NOCO for person answering modules 4, 5, and 6: _____

NOCO for person answering module 7: _____

Check here if Continuation Sheet Used: ☐ Circle number of Continuation Sheets Used: 2 3 4 5 6

Relationship Codes*:			
01. Self	09. Mother	17. Sister-in-law	25. Nephew
02. Husband	10. Older brother	18. Uncle	26. Other
03. Wife	11. Younger Brother	19. Aunt	27. Not Related
04. Son	12. Older Sister	20. Uncle of Husband/Wife	28. Not known
05. Daughter	13. Younger Sister	21. Aunt of Husband/Wife	29. Adopted (with other code)
06. Grandmother	14. Father-in-law	22. Male Cousin	
07. Grandfather	15. Mother-in-law	23. Female Cousin	
08. Father	16. Brother-in-law	24. Niece	

CONTINUE TO INTERVIEW THE HEAD OF HOUSEHOLD

SECTION 3: DISPLACEMENT REGISTER, PAGE 1 CODE OF THIS RESPONDENT (NOCO): ☐☐

I would like to as you some questions about the location of your family of the last few decades – namely between 1974 and 1999. **(INTERVIEWER: IF PERSON ANSERS THAT IS NOT THE HEAD OF HOUSEHOLD, ASK THEM TO REFER TO MOVEMENTS OF HEAD OF HOUSEHOLD.)**

3.1.1 *Where did you live in 1974, the year before the Indonesians came?* (FIRT) Subdistrict _____☐☐ District_____☐☐

3.1.2 *For how long did you live there?* Days Months ☐☐ Years ☐☐ **Exact Date?** Yes ☐ No ☐

3.1.1 *Were you born before or after 1974, the year that the Indonesians came?* Before ☐ After ☐

IF BORN BEFORE 1974, GO TO 3.1.7.

3.1.2 *When were you born?*

Exact Date? Yes ☐ No ☐ **DK** ☐ **Date (mm yyyy):** ☐☐ ☐☐☐☐

Exact Year? Yes ☐ No ☐ **DK** ☐ **Year:** ☐☐☐☐ **Approximation Description:**_____

Exact Month? Yes ☐ No ☐ **DK** ☐ **Month:** ☐☐ **Approximation Description:**_____

3.1.3 *Where were you born?* Suco _____☐☐ Subdistritct _____☐☐ Distritct_____☐☐

Forest? Yes☐No☐**DK**☐ **Refugee Camp?** Yes☐ No☐ **DK**☐ **Stayed with friends/family?** Yes☐No☐**DK**☐ **Moved by the authorities?** Yes☐No☐**DK**☐

3.1.4 *How long did you live there?* Days ☐☐ Months ☐☐ Years ☐☐ **Exact?** Yes ☐ No ☐

IF STILL LIVING IN THIS PLACE AFTER 1999, GO TO SECTION 4.

3.1.5 *What is the primary reason you left? (PMRS)* _____☐

3.1.6 *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

STOP, GO TO 3.2.0.

3.1.7 *Where were you living in 1974, the year before the Indonesians came?* (FIRT)

Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes☐ No☐ **DK**☐ **Refugee Camp?** Yes☐ No☐ **DK**☐ **Stayed with friends/family?** Yes☐ No☐ **DK**☐ **Moved by the authorities?** Yes☐ No☐ **DK**☐

3.1.8 *For how long did you live there?* Days ☐☐ Months ☐☐ Years ☐☐ **Exact?** Yes ☐ No ☐

IF THE RESPONDENT LIVED THERE FOR THE ENTIRE PERIOD OF 1974 TO 1999, GO TO SECTION 4.

3.1.9 *What is the primary reason you left? (PMRS)* _____☐

3.1.10 *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

SECTION 3 CONTINUES ON THE NEXT PAGE

SECTION 3: DISPLACEMENT REGISTER, PAGE 2 CODE OF THIS RESPONDENT (NOCO): ☐☐

3.2.0 *When you first left that place, where did you go, when did you leave, why did you leave and who caused you to leave? Please be as precise as possible – with location names and specific times.*

A

3.2.1: LOCATION (LOCT) Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes□ No□ **DK**□ **Refugee Camp?** Yes□ No□ **DK**□ **Stayed with friends/family?** Yes□ No□ **DK**□ **Moved by the authorities?** Yes□ No□ **DK**□

3.2.2: START (SDAT): Exact Date? Yes □ No □ **Data (mm yyyy):** ☐☐ ☐☐☐☐ **Exact year?** Yes □ No □ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes □ No □ **Approximation description:**_____

Exact month? Yes □ No □ **Month:** ☐☐ **Season?** Rainy □ Dry □ **Approximation description:**_____

END: Exact Date? Yes □ No □ **Data (yy mmmm):** ☐☐ ☐☐☐☐ **Exact year?** Yes □ No □ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes □ No □ **Approximation description:**_____

Exact month? Yes □ No □ **Month:** ☐☐ **Season?** Rainy □ Dry □ **Approximation description:**_____

3.2.3: *What is the primary reason you left?* (PMRS) _____☐

3.2.4: *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

B

3.2.1: LOCATION (LOCT) Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes□ No□ **DK**□ **Refugee Camp?** Yes□ No□ **DK**□ **Stayed with friends/family?** Yes□ No□ **DK**□ **Moved by the authorities?** Yes□ No□ **DK**□

3.2.2: START (SDAT): Exact Date? Yes □ No □ **Data (mm yyyy):** ☐☐ ☐☐☐☐ **Exact year?** Yes □ No □ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes □ No □ **Approximation description:**_____

Exact month? Yes □ No □ **Month:** ☐☐ **Season?** Rainy □ Dry □ **Approximation description:**_____

END: Exact Date? Yes □ No □ **Data (yy mmmm):** ☐☐ ☐☐☐☐ **Exact year?** Yes □ No □ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes □ No □ **Approximation description:**_____

Exact month? Yes □ No □ **Month:** ☐☐ **Season?** Rainy □ Dry □ **Approximation description:**_____

3.2.3: *What is the primary reason you left?* (PMRS) _____☐

3.2.4: *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

SECTION 3 CONTINUES ON THE NEXT PAGE

INTERVIEW FORM # _____ CAVR - RETROSPECTIVE MORTALITY SURVEY

SECTION 3: DISPLACEMENT REGISTER, PAGE 3 CODE OF THIS RESPONDENT (NOCO): ☐☐

C

3.2.1: LOCATION (LOCT) Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes☐ No☐ **DK**☐ **Refugee Camp?** Yes☐ No☐ **DK**☐ **Stayed with friends/family?** Yes☐ No☐ **DK**☐ **Moved by the authorities?** Yes☐ No☐ **DK**☐

3.2.2: START (SDAT): Exact Date? Yes ☐ No ☐ **Data (mm yyyy):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

END: Exact Date? Yes ☐ No ☐ **Data (yy mmmm):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

3.2.3: *What is the primary reason you left?* (PMRS) _____☐

3.2.4: *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

D

3.2.1: LOCATION (LOCT) Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes☐ No☐ **DK**☐ **Refugee Camp?** Yes☐ No☐ **DK**☐ **Stayed with friends/family?** Yes☐ No☐ **DK**☐ **Moved by the authorities?** Yes☐ No☐ **DK**☐

3.2.2: START (SDAT): Exact Date? Yes ☐ No ☐ **Data (mm yyyy):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

END: Exact Date? Yes ☐ No ☐ **Data (yy mmmm):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

3.2.3: *What is the primary reason you left?* (PMRS) _____☐

3.2.4: *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

SECTION 3 CONTINUES ON NEXT PAGE

SECTION 3: DISPLACEMENT REGISTER, PAGE 4 CODE OF THIS RESPONDENT (NOCO): ☐☐

E

3.2.1: LOCATION (LOCT) Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes☐ No☐ **DK**☐ **Refugee Camp?** Yes☐ No☐ **DK**☐ **Stayed with friends/family?** Yes☐ No☐ **DK**☐ **Moved by the authorities?** Yes☐ No☐ **DK**☐

3.2.2: START (SDAT): **Exact Date?** Yes ☐ No ☐ **Data (mm yyyy):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

END: Exact Date? Yes ☐ No ☐ **Data (yy mmmm):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

3.2.3: *What is the primary reason you left?* (PMRS) _____☐

3.2.4: *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

F

3.2.1: LOCATION (LOCT) Suco_____☐☐ Subdistrict _____☐☐ District_____☐☐

Forest? Yes☐ No☐ **DK**☐ **Refugee Camp?** Yes☐ No☐ **DK**☐ **Stayed with friends/family?** Yes☐ No☐ **DK**☐ **Moved by the authorities?** Yes☐ No☐ **DK**☐

3.2.2: START (SDAT): **Exact Date?** Yes ☐ No ☐ **Data (mm yyyy):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

END: Exact Date? Yes ☐ No ☐ **Data (yy mmmm):** ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐

After Indonesians came (12 /1975)? Yes ☐ No ☐ **Approximation description:**_____

Exact month? Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____

3.2.3: *What is the primary reason you left?* (PMRS) _____☐

3.2.4: *Who caused you to leave? (List all reasons that apply.)* (WCLV) _____☐ _____☐

SECTION 3 CONTINUES ON NEXT PAGE

SECTION 3: DISPLACEMENT REGISTER, PAGE 5 CODE OF THIS RESPONDENT (NOCO): ☐☐

Now I would like to ask a few more questions about the places you have lived.

	Yes	No	DK
3.2.5 At any time between *1974 and 1999, were you living in the forest?* IF YES, GO TO PREVIOUS SHEET AND LIST THIS PLACE, THEN GO TO NEXT QUESTION.	☐	☐	☐
3.2.6. *At any time between 1974 and 1999, were you living in a refugee camp?* IF YES, GO TO PREVIOUS SHEET AND LIST THIS PLACE, THEN GO TO NEXT QUESTION.	☐	☐	☐
3.2.7. *At any time between 1974 and 1999, were you staying with family or friends for more than one day?* IF YES, GO TO PREVIOUS SHEET AND LIST THIS PLACE, THEN GO TO NEXT QUESTION.	☐	☐	☐
3.2.8. *At any time between 1974 and 1999, were you assigned a place to live by the military?* IF YES, GO TO PREVIOUS SHEET AND LIST THIS PLACE, THEN GO TO NEXT QUESTION.	☐	☐	☐

Primary Reason Code [3.2.3]
1. Food Shortage/Famine-related
2. Afraid of authorities
2. Authorities forced you to move
4. Family reunion
5. Move to city for work/other opportunity
6. Other (List Reason)
7. Wanted more farming/garden space
8. Don't know

CAUSE OF LEAVE CODES [3.2.4]
A. Indonesian Military [State Division]
B. Police [State Police Unit Name]
C. Militia [State Militia Group]
D. Intelligence [State Name of Group]
E. Political Party [State Political Party]
F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih]
G. Falinitil
H. Unknown Military Force
J. Unknown
K. Other [List Name/Description]

SECTION 4: ADULT PARENTAL SURVIVAL, PAGE 1

INTERVIEWER: PICK RANDOM ADULT TO ANSWER NEXT THREE SECTIONS.

Thank you. Now I would like to speak with [insert name here] for a while. We will talk about other members of [name] family.

CODE OF THIS RESPONDENT (NOCO): ☐☐

	(INTERVIEWER: THIS COLUMN FIRST)	CODE		(INTERVIEWER: THIS COLUMN SECOND)	CODE
4.1.2	**INTERVIEWER: IF THE RESPONDENT'S MOTHER IS LISTED IN SECTION 2, USE THE NOCO (CODE) FROM THERE, OTHERWISE USE NOCO 42:** **NOCO** ☐☐		4.2.2	**INTERVIEWER: IF THE RESPONDENT'S FATHER IS LISTED IN SECTION 2, USE THE NOCO (CODE) FROM THERE, OTHERWISE USE NOCO 42:** **NOCO** ☐☐	
4.1.3	*What is your mother's full name?*		4.2.3	*What is your father's full name?*	
(a)	First name (MFNM)_____		(a)	First name: (FFNM)_____	
(b)	Last Name (MLNM)_____		(b)	Last Name (FLNM)_____	
(c)	Nickname (MNNM)_____		(c)	Nickname (FNNM):_____	
4.1.4	*When was your mother born?* (MBIR) **Exact Date?** Yes ☐ No ☐ **Date(dmy):** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____		4.2.4	*When was your father born?* (FBIR) **Exact Date?** Yes ☐ No ☐ **Date(dmy):** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____	
4.1.5	*Where was she born?*		4.2.5	*Where was he born?*	
(a)	Suco (MBSU)_____	☐☐	(a)	Suco (FBSU)_____	☐☐
(b)	Subdistrict (MBSB)_____	☐☐	(b)	Subdistrict (FBSB)_____	☐☐
(c)	District (MBDI)_____	☐☐	(c)	District (FBDI)_____	☐☐

SECTION 4 CONTINUES ON NEXT PAGE

	(INTERVIEWER: THIS COLUMN FIRST)	CODE		(INTERVIEWER: THIS COLUMN SECOND)	CODE
4.1.6	*Is your mother alive, dead, or disappeared?* Alive=1, Dead=2, Disappeared=3, Don't Know=4 (MOST)	□ 1 → STOP 2 →4.1.10 3/4 → 4.1.7	4.2.6	*Is your father alive, dead, or disappeared?* Alive=1, Dead=2, Disappeared=3, Don't Know=4 (FAST)	□ 1 → STOP 2 → 4.2.10 3/4 → 4.2.7
4.1.7	*When was the last time you had contact with your mother?*(MLCT) **Exact Date?** Yes □ No □ **Date(dmy):** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**_____ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**_____		4.2.7	*When was the last time you had contact with your father?* (FLCT) **Exact Date?** Yes □ No □ **Date(dmy):** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**_____ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**_____	
4.1.8	*Where did you last have contact with your mother?*		4.2.8	*Where did you last have contact with your father?*	
(a)	Suco (MLCS) _____	□□	(a)	Suco (MLCS) _____	□□
(b)	Subdistrict (MLCU) _____	□□	(b)	Subdistrict (MLCU) _____	□□
(c)	District (MLCD) _____ **Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	□□ STOP	(c)	District (MLCD) _____ **Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	□□ STOP
4.1.9	*If you mother disappeared, how did she disappear?*	□□ STOP	4.2.9	*If your father disappeared, how did he disappear?*	□□ STOP
4.1.10	*When did your mother die?* (MDOD) **Exact Date?** Yes □ No □ **Date(dmy):** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**_____ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**_____		4.2.10	*When did your father die?* (FDOD) **Exact Date?** Yes □ No □ **Date(dmy):** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**_____ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**_____	

SECTION 4 CONTINUES ON NEXT PAGE

SECTION 4: ADULT PARENTAL SURVIVAL, PAGE 3 CODE OF THIS RESPONDENT (NOCO): ☐☐

4.1.11	*Where did your mother die?*		4.2.11	*Where did your father die?*	
	(a) Suco (MDDS) _____	☐☐		(a) Suco (MDDS) _____	☐☐
	(b) Subdistrict (MDDU) _____	☐☐		(b) Subdistrict (MDDU) _____	☐☐
	(c) District (MDDD) _____ **Forest?**Yes☐ No☐DK☐ **Refugee Camp?**Yes☐ No☐ DK☐	☐☐		(c) District (MDDD) _____ **Forest?**Yes☐ No☐DK☐ **Refugee Camp?**Yes☐ No☐ DK☐	☐☐
4.1.12	*How did your mother die?* (CMTD)	☐☐ ☐	4.2.12	*How did your father die?* (CFTD)	☐☐ ☐
4.1.13	*Was your mother either buried in a grave or was there a memorial erected for her or both?* (MGRI) Yes=1, No=2, Don't Know=3 **Bones in Forest?** Yes☐ No☐	☐ 1 → 4.1.13 2/3 → STOP	4.2.13	*Was your father either buried in a grave or was there a memorial erected for her or both?* (FGRI) Yes=1, No=2, Don't Know=3 **Bones in Forest?** Yes☐ No☐	☐ 1 → 4.2.13 2/3 → STOP
4.1.14	*Where is the graveyard/memorial located?*		4.2.14	*Where is the graveyard/memorial located?*	
	(a) Public cemetery ☐ Private cemetery ☐ (FGYT)			(a) Public cemetery ☐ Private cemetery ☐ (FGYT)	
	(b) Cemetery_____ (FGYG)			(b) Cemetery_____ (FGYG)	
	(c) Suco _____☐☐ (FGYS)			(c) Suco _____☐☐ (FGYS)	
	(d) Subdistrict _____☐☐ (FGYU)			(d) Subdistrict _____☐☐ (FGYU)	
	(e) District _____☐☐ (FGYD)			(e) District _____☐☐ (FGYD)	
4.1.15	*Is her name written on the grave/memorial?* Yes=1, No=2, DK=3 (MGYN)	☐	4.2.15	*Is his name written on the grave/memorial?* Yes=1, No=2, DK=3 (FGYN)	☐
4.1.16	*What material is the grave/memorial made from?* (MGYM)		4.2.16	*What material is the grave/memorial made from?* (FGYM)	
	Stone = 1, Cement = 2, Wood = 3, Dirt = 4, Don't Know = 5, Other = 6 _____	☐		Stone = 1, Cement = 2, Wood = 3, Don't Know = 5 , Other = 6 _____	☐
	(INTERVIEWER: GO TO NEXT COLUMN)			**(INTERVIEWER: GO TO NEXT SECTION)**	

SECTION 4 CONTINUES ON NEXT PAGE

CAUSE OF DEATH CODES [4.1.9/4.2.9/4.1.12/4.2.12]

Extrajudicial execution	1	Sick and Hungry	7
Civilian death in conflict	2	Sick and not hungry	8
Combatant death in conflict	3	Death in childbirth	9
Death due to torture	4	Natural Causes	10
Death due to mistreatment by authorities	5	Other (_____)	11
Hunger/starvation	6	Don't know	12

SECURITY CODE [4.1.9/4.2.9/4.1.12/4.2.12]

A. Indonesian Military [State Division]	F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih]
B. Police [State Police Unit Name]	G. Falinitil
C. Militia [State Militia Group]	H. Unknown Military Force
D. Intelligence [State Name of Group]	J. Unknown
E. Political Party [State Political Party]	K. Other [List Name/Description]

SECTION 5: ADULT SIBLINGS, PAGE 1

Thank you. Now I will ask you, one person at a time, about your older brothers, older sisters, and younger brothers and sisters.

NOCO FOR THIS RESPONDENT (NOCO): □□

No.	QUESTIONS AND FILTERS				CODE
5.2	*Now, I would like to ask you some questions about your brothers and sisters, that is the children who were born to your natural mother, including those who are living with you, those who are living elsewhere and those who have died. How many children has your mother ever had?* (MNPC)				□□ "01" or "Only One"→5.17
		(1)	**(2)**	**(3)**	**(4)**
5.3	**INTERVIEWER: IF RESPONDENT'S SIBLING IS LISTED IN SECTION 2, USE THE NOCO (CODE) FROM THERE.**	□□ (SECTION 2 NOCO OR USE NOCO 51)	□□ (SECTION 2 NOCO OR USE NOCO 52)	□□ (SECTION 2 NOCO OR USE NOCO 53)	□□ (SECTION 2 NOCO OR USE NOCO 54)
5.4	*What is the name given to your oldest (next oldest) sibling? (First, Last, Nickname)* (SFNM, SLNM, SNNM)	F_____ L_____ N_____	F_____ L_____ N_____	F_____ L_____ N_____	F_____ L_____ N_____
5.5	*Is (NAME) male or female?* (SGND) M = 1, F = 2	□	□	□	□
5.6	*Is (NAME) alive, dead, or disappeared?* Alive=1, Dead=2, Disappeared=3, Don't Know=4 (SSAL)	□ 1 → (2) 2 →5.10 3/4 → 5.7	□ 1 → (3) 2 →5.10 3/4 → 5.7	□ 1 → (4) 2 →5.10 3/4 → 5.7	□ 1 → (5) 2 →5.10 3/4 → 5.7
5.7	*Where is the last place you had contact with [NAME]?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	Suco (SLCS)	_____□□	_____□□	_____□□	_____□□
(b)	Subdistrict (SLCU)	_____□□	_____□□	_____□□	_____□□
(c)	District (SLCD)	_____□□	_____□□	_____□□	_____□□

SECTION 5: ADULT SIBLINGS, PAGE 2 NOCO FOR THIS RESPONDENT (NOCO): □□

5.8	*When was the last time you had contact with him/her?* (SLCT)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________
5.9	*If [NAME] disappeared, how did he disappear?*	□□ □ **STOP → (2)**	□□ □ **STOP → (3)**	□□ □ **STOP → (4)**	□□ □ **STOP → (5)**
5.10	*Where did he/she die?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	Suco (SDDS)	______________□□	______________□□	______________□□	______________□□
(b)	Subdistrict (SDDU)	______________□□	______________□□	______________□□	______________□□
(c)	District (SDDD)	______________□□	______________□□	______________□□	______________□□

SECTION 5 CONTINUES ON NEXT PAGE

		(1)	(2)	(3)	(4)
5.11	*When did he/she die?* (dd mm yyyy) (SDOD)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________
5.12	How did he/she die? Cause of Death, Perpetrator (SCDC), (SPDC)	□□ □	□□ □	□□ □	□□ □

CAUSE OF DEATH CODES [5.8/5.12]

Extrajudicial execution	1	Sick and Hungry	7
Civilian death in conflict	2	Sick and not hungry	8
Combatant death in conflict	3	Death in childbirth	9
Death due to torture	4	Natural Causes	10
Death due to mistreatment by authorities	5	Other (____________________)	11
Hunger/starvation	6	Don't know	12

SECURITY CODE [5.8/5.12]

A. Indonesian Military [State Division]	F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih]
B. Police [State Police Unit Name]	G. Falinitil
C. Militia [State Militia Group]	H. Unknown Military Force
D. Intelligence [State Name of Group]	J. Unknown
E. Political Party [State Political Party]	K. Other [List Name/Description]

SECTION 5 CONTINUES ON NEXT PAGE

SECTION 5: ADULT SIBLINGS, PAGE 4 NOCO FOR THIS RESPONDENT (NOCO): ☐☐

		(1)	(2)	(3)	(4)
5.13	*Was s/he buried in a grave?* Yes=1, No=2, Don't Know=3 (SGRG)	☐ 2/3 → (2) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (3) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (4) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (5) **Bones in Forest?** Yes☐ No☐ DK☐
5.14	Where is the grave located? (SGRP) Public cemetery = 1, Private cemetery = 2	☐	☐	☐	☐
	(SGRV) cemetery	_____	_____	_____	_____
	(SGRS) Suco	_____☐☐	_____☐☐	_____☐☐	_____☐☐
	(SGRU) Subdistrict	_____☐☐	_____☐☐	_____☐☐	_____☐☐
	(SGRD) District	_____☐☐	_____☐☐	_____☐☐	_____☐☐
5.15	Is his/her name on the monument? *Yes=1, No=2, LH=3 (SGYN)*	☐	☐	☐	☐
5.16	*What is the grave made from?* (SGRM) Stone = 1, Cement = 2, Wood = 3, Dirt = 4, Don't Know = 5, Other = 6	☐ _____	☐ _____	☐ _____	☐ _____

Check here if Continuation Sheet Used: ☐

Circle number of Continuation Sheets Used: 1 2 3 4 5 6

NOCO FOR THIS RESPONDENT (NOCO): ☐☐

		(5)	**(6)**	**(7)**	**(8)**
5.3	**INTERVIEWER: IF RESPONDENT'S SIBLING IS LISTED IN SECTION 2, USE THE NOCO (CODE) FROM THERE.**	☐☐ (SECTION 2 NOCO OR USE NOCO 55)	☐☐ (SECTION 2 NOCO OR USE NOCO 56)	☐☐ (SECTION 2 NOCO OR USE NOCO 57)	☐☐ (SECTION 2 NOCO OR USE NOCO 58)
5.4	*What is the name given to your oldest (next oldest) sibling? (First, Last Name, Nickname)* (SFNM, SLNM, SNNM)	F_____ L_____ N_____	F_____ L_____ N_____	F_____ L_____ N_____	F_____ L_____ N_____
5.5	*Is (NAME) male or female?* (SGND) M = 1, F = 2	☐	☐	☐	☐
5.6	*Is (NAME) alive, dead, or disappeared?* Alive=1, Dead=2, Disappeared=3, Don't Know=4 (SSAL)	☐ 1 → (6) 2 →5.10 3/4 → 5.7	☐ 1 → (7) 2 →5.10 3/4 → 5.7	☐ 1 → (8) 2 →5.10 3/4 → 5.7	☐ 1 → (9) 2 →5.10 3/4 → 5.7
5.7	*Where is the last place you had contact with [NAME]?*	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐
(a)	Suco (SLCS)	_____☐☐	_____☐☐	_____☐☐	_____☐☐
(b)	Subdistrict (SLCU)	_____☐☐	_____☐☐	_____☐☐	_____☐☐
(c)	District (SLCD)	_____☐☐	_____☐☐	_____☐☐	_____☐☐

SECTION 5 CONTINUES ON NEXT PAGE

SECTION 5: ADULT SIBLINGS, PAGE 2, CONTINUATION # 1 2 3 4

NOCO FOR THIS RESPONDENT (NOCO): □□

		(5)	(6)	(7)	(8)
5.8	*When was the last time you had contact with him/her?* (SLCT)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________
5.9	*If [NAME] disappeared, how did he disappear?*	□□ □ **STOP → (6)**	□□ □ **STOP → (7)**	□□ □ **STOP → (8)**	□□ □ **STOP → (9)**
5.10	*Where did he/she die?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	Suco (SDDS)	______________□□	______________□□	______________□□	______________□□
(b)	Subdistrict (SDDU)	______________□□	______________□□	______________□□	______________□□
(c)	District (SDDD)	______________□□	______________□□	______________□□	______________□□

SECTION 5 CONTINUES ON NEXT PAGE

NOCO FOR THIS RESPONDENT (NOCO): ☐☐

		(5)	(6)	(7)	(8)
5.11	*When did he/she die?* (dd mm yyyy) (SDOD)	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________
5.12	How did he/she die? Cause of Death, Perpetrator (SCDC), (SPDC)	☐☐ ☐	☐☐ ☐	☐☐ ☐	☐☐ ☐

CAUSE OF DEATH CODES [5.8/5.12]

Extrajudicial execution	1	Sick and Hungry	7
Civilian death in conflict	2	Sick and not hungry	8
Combatant death in conflict	3	Death in childbirth	9
Death due to torture	4	Natural Causes	10
Death due to mistreatment by authorities	5	Other (____________________)	11
Hunger/starvation	6	Don't know	12

SECURITY CODE [5.8/5.12]

A. Indonesian Military [State Division]	F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih]
B. Police [State Police Unit Name]	G. Falinitil
C. Militia [State Militia Group]	H. Unknown Military Force
D. Intelligence [State Name of Group]	J. Unknown
E. Political Party [State Political Party]	K. Other [List Name/Description]

SECTION 5 CONTINUES ON NEXT PAGE

		(5)	(6)	(7)	(8)
5.13	*Was s/he buried in a grave?* Yes=1, No=2, Don't Know=3 (SGRG)	☐ 2/3 → (6) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (7) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (8) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (9) **Bones in Forest?** Yes☐ No☐ DK☐
5.14	Where is the grave located? (SGRP) Public cemetery = 1, Private cemetery = 2 (SGRV) cemetery (SGRS) Suco (SGRU) Subdistrict (SGRD) District	☐ _____ _____☐☐ _____☐☐ _____☐☐	☐ _____ _____☐☐ _____☐☐ _____☐☐	☐ _____ _____☐☐ _____☐☐ _____☐☐	☐ _____ _____☐☐ _____☐☐ _____☐☐
5.15	Is his/her name on the monument? *Yes=1, No=2, LH=3 (SGYN)*	☐	☐	☐	☐
5.16	*What is the grave made from?* (SGRM) Stone = 1, Cement = 2, Wood = 3, Dirt = 4, Don't Know = 5, Other = 6	☐ _____	☐ _____	☐ _____	☐ _____

If Additional Continuation Sheet Used, Check Here: ☐

SECTION 6: ADULT HUMAN RIGHTS HISTORY PAGE 1

I have one more topic to ask you about. We have heard and are concerned about the possibility of human rights abuses in East Timor, including torture, imprisonment, beating, and property destruction, and especially sexual assault and rape of East Timorese women by any military forces. We acknowledge that human rights abuses - whether they were committed recently or a long time ago - are very painful and therefore difficult to talk about. What human rights abuses have been experienced by any of your family members or yourself?

	Yes	No	DK
6.1. *Have you or another family member been disappeared or separated from the rest of the family?*	☐	☐	☐
6.2. *Have you or another family member been imprisoned?*	☐	☐	☐
6.3. *Have you or another family member been beaten?*	☐	☐	☐
6.4. *Have you or another family member been tortured?*	☐	☐	☐
6.5. *Has a member of your family been killed?*	☐	☐	☐
6.6. *Have you or another family member suffered a gunshot wound?*	☐	☐	☐
6.7. *Have you or another family member been assaulted sexually?*	☐	☐	☐
6.8. *Have you or another family member been raped?*	☐	☐	☐
6.9. *Have you or another family member had property destroyed?*	☐	☐	☐
6.10. *Have you or another family member been displaced by the authorities (e.g., children to Java)?*	☐	☐	☐
6.11. *Have you or another family member experienced another human rights violation?*	☐	☐	☐

INTERVIEWER: IF RESPONDENT ANSWERS YES TO ANY OF THESE QUESTIONS, FILL OUT DETAILS IN THE CHART.

INTERVIEW FORM # ____________ CAVR - RETROSPECTIVE MORTALITY SURVEY

SECTION 6: ADULT HUMAN RIGHTS HISTORY PAGE 2 CODE OF THIS RESPONDENT (NOCO): ☐☐

Person (NOCO)	Abuse (HRAB)	Where abuse or death occurred (HRWH)	Month and year of abuse or death (HRYR)	Perpetrator (HRPR)	Witness the abuse or after-effects (HRWT)	If death, cause of death (HRDC)
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐____________	☐	☐☐
NOCO	1=sep. & disap. 2=imprisonment 3=beating 4=torture 5=killing 6=gunshot wound 7=sexual assault, no rape 8=rape 9=property destruction 10=other (specify) 88=no response	1=home 2=hospital 3=police station 4=prison 5=torture center 6=police station 7=streets 8=work 9=other (specify) 99=DK 88=NR	Write-in	A=Indonesian Military (Specify Division) B=Police (Specify Police Unit Name) C=Militia (State group) D=Intelligence (State group) E=Political party (State party) F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih] G=Falintil H. Unknown Military Force J. Unknown K. Other [List Name/Description] 88=NR	1=witnessed abuse 2=after effects	*1=extrajudicial execution* *2=civilian death in conflict* 3=combatant death in conflict 4=death due to torture 5=death due to mistreatment by authorities 6=hunger/starvation 7=sick and hungry 8=sick and not hungry 9=death in childbirth 10=natural causes 11=other 12=don't know

Check here if continuation sheet used: ☐

Circle number of Continuation Sheets Used: 2 3 4 5 6

INTERVIEW FORM # ____________ CAVR - RETROSPECTIVE MORTALITY SURVEY

SECTION 6: ADULT HUMAN RIGHTS HISTORY CONTINUATION PAGE # 1 2 3 4 5 6

CODE OF THIS RESPONDENT (NOCO): ☐☐

Person (NOCO)	Abuse (HRAB)	Where abuse or death occurred (HRWH)	Month and year of abuse or death (HRYR)	Perpetrator (HRPR)	Witness the abuse or after-effects (HRWT)	If death, cause of death (HRDC)
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
☐☐	☐☐	☐☐	☐☐ ☐☐☐☐	☐☐ ____________	☐	☐☐
NOCO	1=sep. & disap. 2=imprisonment 3=beating 4=torture 5=killing 6=gunshot wound 7=sexual assault, no rape 8=rape 9=property destruction 10=other (specify) 88=no response	1=home 2=hospital 3=police station 4=prison 5=torture center 6=police station 7=streets 8=work 9=other (specify) 99=DK 88=NR	Write-in	A=Indonesian Military (Specify Division) B=Police (Specify Police Unit Name) C=Militia (State group) D=Intelligence (State group) E=Political party (State party) F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih] G=Falintil H. Unknown Military Force J. Unknown K. Other [List Name/Description] 88=NR	1=witnessed abuse 2=after effects	*1=extrajudicial execution* *2=civilian death in conflict* 3=combatant death in conflict 4=death due to torture 5=death due to mistreatment by authorities 6=hunger/starvation 7=sick and hungry 8=sick and not hungry 9=death in childbirth 10=natural causes 11=other 12=don't know

Check here if additional continuation sheet used: ☐

SECTION 7: BIRTH HISTORY, PAGE 1 – **INTERVIEWER: PICK RANDOM ADULT FEMALE TO ANSWER THIS PART.**

For this part, we would like to speak with [NAME] in private. NOCO FOR THIS RESPONDENT (NOCO): ☐☐

No.	QUESTIONS AND FILTERS				CODE
	Now, I would like to ask you some questions about your pregnancy history – that is all your pregnancies (including abortions, miscarriages and still-births) and the children who were born to you, including those who are living with you, those who are living elsewhere and those who have died				
7.1	*How many pregnancies have you ever had (including live births, still-births, miscarriages and abortions)?* (NOPR)				☐☐ If response is "0" → STOP
7.2	*How many live births have you ever had?* (NOLB)				☐☐ If response is "0 "→ STOP
		(1)	**(2)**	**(3)**	**(4)**
7.3	**INTERVIEWER: IF RESPONDENT'S SIBLING IS LISTED IN SECTION 2, USE THE NOCO (CODE) FROM THERE.**	☐☐ (SECTION 2 NOCO OR USE NOCO 71)	☐☐ (SECTION 2 NOCO OR USE NOCO 72)	☐☐ (SECTION 2 NOCO OR USE NOCO 73)	☐☐ (SECTION 2 NOCO OR USE NOCO 74)
7.4	*What is the name given to your oldest (next oldest) son or daughter?* **F**irst/**L**ast/**N**ickname (CFNM, CLMN, CNNM)	F_____ L_____ N_____	F_____ L_____ N_____	F_____ L_____ N_____	F_____ L_____ N_____
7.5	*Is (NAME) male or female?* M = 1, F = 2 (CGND)	☐	☐	☐	☐

SECTION 7 CONTINUES ON NEXT PAGE

NOCO FOR THIS RESPONDENT (NOCO): □□

		(1)	(2)	(3)	(4)
7.6	*When was (NAME) born?* (dd mm yyyy) (CBIR)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________
7.7	*Is [NAME] Alive, Dead, Disappeared?* Alive=1, Dead=2, Disappeared=3, Don't Know=4 (CLIV)	□ 1 → (2) 2 → 7.10 3/4 → 7.7	□ 1 → (3) 2 → 7.10 3/4 → 7.7	□ 1 → (4) 2 → 7.10 3/4 → 7.7	□ 1 → (5) 2 → 7.10 3/4 → 7.7
7.8	*Where was the last place you had contact with [NAME]?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	*Suco (CLCS)*	______________□□	______________□□	______________□□	______________□□
(b)	Subdistrict (CLCU)	______________□□	______________□□	______________□□	______________□□
(c)	District (CLCD)	______________□□	______________□□	______________□□	______________□□

SECTION 7 CONTINUES ON NEXT PAGE

SECTION 7: BIRTH HISTORY, PAGE 3 NOCO FOR THIS RESPONDENT (NOCO): □□

		(1)	(2)	(3)	(4)
7.9	*When was the last time you had contact with (NAME)?* (dd mm yyyy) (CLCT)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________
7.10	*If [NAME] disappeared, how did he/she disappear?* Cause of Death Code (CCDC), Perpetrator Code (CPDC)	□□ □ **STOP → (2)**	□□ □ **STOP → (3)**	□□ □ **STOP → (4)**	□□ □ **STOP → (5)**
7.11	*Where did he/she die?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	Suco (CDDS)	______________□□	______________□□	______________□□	______________□□
(b)	Subdistrict (CDDU)	______________□□	______________□□	______________□□	______________□□
(c)	District (CDDD)	______________□□	______________□□	______________□□	______________□□

SECTION 7 CONTINUES ON NEXT PAGE

SECTION 7: BIRTH HISTORY, PAGE 4 NOCO FOR THIS RESPONDENT (NOCO): ☐☐

		(1)	(2)	(3)	(4)
7.12	*When did he/she die?* (CDOD)	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________	**Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**______________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**______________
7.13	*How did he/she die?* Cause of Death Code (CCDC), Perpetrator Code (CPDC)	☐☐ ☐	☐☐ ☐	☐☐ ☐	☐☐ ☐
7.14	*Was s/he buried in a grave?* Yes=1, No=2, Don't Know=3 (CGRG)	☐ 2/3 → (2) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (3) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (4) **Bones in Forest?** Yes☐ No☐ DK☐	☐ 2/3 → (5) **Bones in Forest?** Yes☐ No☐ DK☐

SECTION 7 CONTINUES ON NEXT PAGE

		(1)	(2)	(3)	(4)
7.15	Where is the grave located? (SGRP) Public cemetery = 1, Private cemetery = 2	☐	☐	☐	☐
	(CGRV) cemetery	_____	_____	_____	_____
	(CGRS) Suco	_____ ☐☐	_____ ☐☐	_____ ☐☐	_____ ☐☐
	(CGRU) Subdistrict	_____ ☐☐	_____ ☐☐	_____ ☐☐	_____ ☐☐
	(CGRD) District	_____ ☐☐	_____ ☐☐	_____ ☐☐	_____ ☐☐
7.16	Is his/her name on the monument? *Yes=1, No=2, LH=3 (SGYN)*	☐	☐	☐	☐
7.17	*What was the grave made from?* (CGRM) Stone = 1, Cement = *2, Wood = 3, Dirt =4, Don't Know = 5, Other = 6*	☐ _____	☐ _____	☐ _____	☐ _____

CAUSE OF DEATH CODES [7.9/7.12]

Extrajudicial execution	1	Sick and Hungry	7
Civilian death in conflict	2	Sick and not hungry	8
Combatant death in conflict	3	Death in childbirth	9
Death due to torture	4	Natural Causes	10
Death due to mistreatment by authorities	5	Other (_____)	11
Hunger/starvation	6	Don't know	12

SECURITY CODE [5.8/5.12]

A. Indonesian Military [State Division]	F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih]
B. Police [State Police Unit Name]	G. Falinitil
C. Militia [State Militia Group]	H. Unknown Military Force
D. Intelligence [State Name of Group]	J. Unknown
E. Political Party [State Political Party]	K. Other [List Name/Description]

Check here if continuation sheet used: ☐ Circle number of Continuation Sheets Used: 2 3 4 5 6

SECTION 7: BIRTH HISTORY, CONTINUATION PAGE 1

NOCO FOR THIS RESPONDENT (NOCO): □□

		(5)	(6)	(7)	(8)
7.3	**INTERVIEWER: IF RESPONDENT'S SIBLING IS LISTED IN SECTION 2, USE THE NOCO (CODE) FROM THERE.**	□□ (SECTION 2 NOCO OR USE NOCO 75)	□□ (SECTION 2 NOCO OR USE NOCO 76)	□□ (SECTION 2 NOCO OR USE NOCO 77)	□□ (SECTION 2 NOCO OR USE NOCO 78)
7.4	*What is the name given to your oldest (next oldest) son or daughter?* **F**irst/**L**ast/**N**ickname (CFNM, CLMN, CNNM)	F________ L________ N________	F________ L________ N________	F________ L________ N________	F________ L________ N________
7.5	*Is (NAME) male or female?* M = 1, F = 2 (CGND)	□	□	□	□
7.6	*When was (NAME) born?* (dd mm yyyy) (CBIR)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**________

SECTION 7 CONTINUES ON NEXT PAGE

SECTION 7: BIRTH HISTORY, CONTINUATION PAGE 2 NOCO FOR THIS RESPONDENT (NOCO): □□

		(5)	(6)	(7)	(8)
7.7	*Is [NAME] Alive, Dead, Disappeared?* Alive=1, Dead=2, Disappeared=3, Don't Know=4 (CLIV)	□ 1 → (6) 2 /3 → 7.10 4 → 7.7	□ 1 → (7) 2 /3 → 7.10 4 → 7.7	□ 1 → (8) 2 /3 → 7.10 4 → 7.7	□ 1 → (9) 2/3 → 7.10 4 → 7.7
7.8	*Where was the last place you had contact with [N]?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	Suco (CLCS)	______________□□	______________□□	______________□□	______________□□
(b)	Subdistrict (CLCU)	______________□□	______________□□	______________□□	______________□□
(c)	District (CLCD)	______________□□	______________□□	______________□□	______________□□
7.9	*When was the last time you had contact with (NAME)?* (dd mm yyyy) (CLCT)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________

SECTION 7 CONTINUES ON NEXT PAGE

SECTION 7: BIRTH HISTORY, CONTINUATION PAGE 3 NOCO FOR THIS RESPONDENT (NOCO): □□

		(5)	(6)	(7)	(8)
7.10	*If [NAME] disappeared, how did he/she disappear?* Cause of Death Code (CCDC), Perpetrator Code (CPDC)	□□ □ **STOP → (6)**	□□ □ **STOP → (7)**	□□ □ **STOP → (8)**	□□ □ **STOP → (9)**
7.11	*Where did he/she die?*	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□
(a)	Suco (CDDS)	______________□□	______________□□	______________□□	______________□□
(b)	Subdistrict (CDDU)	______________□□	______________□□	______________□□	______________□□
(c)	District (CDDD)	______________□□	______________□□	______________□□	______________□□
7.12	*When did he/she die?* (CDOD)	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________	**Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______________ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______________

SECTION 7 CONTINUES ON NEXT PAGE

INTERVIEW FORM # _____ CAVR - RETROSPECTIVE MORTALITY SURVEY

SECTION 7: BIRTH HISTORY, CONTINUATION PAGE 4 NOCO FOR THIS RESPONDENT (NOCO): ☐☐

		(5)	(6)	(7)	(8)
7.13	*How did he/she die?* Cause of Death Code (CCDC), Perpetrator Code (CPDC)	☐☐ ☐	☐☐ ☐	☐☐ ☐	☐☐ ☐
7.14	*Was he/she buried in a grave?* Yes=1, No=2, Don't Know=3 (CGRV)	☐ 2/3 → (6) **Bones in Forest?** Yes□ No□ DK□	☐ 2/3 → (7) **Bones in Forest?** Yes□ No□ DK□	☐ 2/3 → (8) **Bones in Forest?** Yes□ No□ DK□	☐ 2/3 → (9) **Bones in Forest?** Yes□ No□ DK□
7.15	Where is the grave located? (SGRP) Public cemetery = 1, Private cemetery = 2	☐	☐	☐	☐
	(CGRV) cemetery	_____	_____	_____	_____
	(CGRS) Suco	_____☐☐	_____☐☐	_____☐☐	_____☐☐
	(CGRU) Subdistrict	_____☐☐	_____☐☐	_____☐☐	_____☐☐
	(CGRD) District	_____☐☐	_____☐☐	_____☐☐	_____☐☐
7.16	Is his/her name on the monument? *Yes=1, No=2, LH=3 (SGYN)*	☐	☐	☐	☐
7.17	*What was the grave made from? (CGRM) Stone=1, Cement=2, Wood = 3, Dirt=4, DK=5, Other = 6*	☐ _____	☐ _____	☐ _____	☐ _____

SECTION 7 CONTINUES ON NEXT PAGE

SECTION 7: BIRTH HISTORY, CONTINUATION PAGE 5 NOCO FOR THIS RESPONDENT (NOCO): ☐☐

CAUSE OF DEATH CODES [7.9/7.12]

Extrajudicial execution	1	Sick and Hungry	7
Civilian death in conflict	2	Sick and not hungry	8
Combatant death in conflict	3	Death in childbirth	9
Death due to torture	4	Natural Causes	10
Death due to mistreatment by authorities	5	Other (_____)	11
Hunger/starvation	6	Don't know	12

SECURITY CODE [5.8/5.12]

A. Indonesian Military [State Division]	F. Civil Defense/Paramilitary [hansip, wanra, kamsa, ratih]
B. Police [State Police Unit Name]	G. Falinitil
C. Militia [State Militia Group]	H. Unknown Military Force
D. Intelligence [State Name of Group]	J. Unknown
E. Political Party [State Political Party]	K. Other [List Name/Description]

Check here if additional continuation sheet used: ☐

7.18	*Did any of your pregnancies result in a miscarriage?* Yes=1, No=2, Don't Know=3 (MCMK)		☐ No/DK → 7.21
7.19	*How many miscarriages have you had?* (MCCT)		☐☐
7.20	*When and where did you have the miscarriage(s)?*		
	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐

SECTION 7 CONTINUED ON NEXT PAGE

7.20	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐
7.21	*Did any of your pregnancies end in abortion? Yes=1, No=2, Don't Know=3 (ABMK)*		☐ ☐ No/DK → 7.24
7.22	*How many abortions have you had?* (ABCT)		☐☐
7.23	*When and where did you have the abortion(s)?*		
	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐

SECTION 7 CONTINUED ON NEXT PAGE I

7.23	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□ ______□□ ______□□ ______□□
	(DMC) **Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□ ______□□ ______□□ ______□□
	(DMC) **Exact Date?** Yes □ No □ **Date:** □□ □□ □□□□ **Exact year?** Yes □ No □ **Year :** □□□□ **After Indonesians came (12 /1975)?** Yes □ No □ **Approximation description:**______ **Approximate Month?** Yes □ No □ **Month:** □□ **Season?** Rainy □ Dry □ **Approximation description:**______	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes□ No□ DK□ **Refugee Camp?** Yes□ No□ DK□ ______□□ ______□□ ______□□

SECTION 7 CONTINUED ON NEXT PAGE

7.23	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐
7.24	*Did any of your pregnancies end in still-birth?* Yes=1, No=2, Don't Know=3 (SBMK)	☐ No/DK → STOP	
7.25	*How many still-births have you had?* (SBCT)	☐☐	
7.25	*When and where did the still-birth(s) occur?*		
	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**_____ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**_____	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ _____☐☐ _____☐☐ _____☐☐

SECTION 7 CONTINUED ON NEXT PAGE

 NOCO FOR THIS RESPONDENT (NOCO): ☐☐

7.25	(a) Date:	(b) Location:	
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**________________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**________________	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ __________☐☐ __________☐☐ __________☐☐
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**________________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**________________	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ __________☐☐ __________☐☐ __________☐☐
	(DMC) **Exact Date?** Yes ☐ No ☐ **Date:** ☐☐ ☐☐ ☐☐☐☐ **Exact year?** Yes ☐ No ☐ **Year :** ☐☐☐☐ **After Indonesians came (12 /1975)?** Yes ☐ No ☐ **Approximation description:**________________ **Approximate Month?** Yes ☐ No ☐ **Month:** ☐☐ **Season?** Rainy ☐ Dry ☐ **Approximation description:**________________	Suco (MCS1) Subdistrict (MCU1) District (MCD1)	**Forest?** Yes☐ No☐ DK☐ **Refugee Camp?** Yes☐ No☐ DK☐ __________☐☐ __________☐☐ __________☐☐

Interview End Time:____________________

Appendix 1 – Abbreviations

Abbreviation	Full Name or Description
ABRI	Angkatan Bersenjata Republik Indonesia (Armed Forces of the Republic of Indonesia), which is now known as Tentara Nasional Indonesia (TNI)
AI	Amnesty International
Apodeti	Associaçao Popular Democrática de Timor Pró-Refrendum (Pro Referendum Popular Democratic Association of Timor)
BPS	Biro Pusat Statistik (Central Statistics Bureau, Indonesia)
CAVR	Commissao de Acolhimento, Verdade e Reconciliacao (Commission for Reception, Truth and Reconciliation)
	Crude Death Rate
CRS	Catholic Relief Services
DOB	Date of Birth
DOD	Date of Death
ENAR	Eastern North American Region, International Biometric Society
Falintil	Forças Armadas de Libertação Nacional de Timor Leste (Armed Forces of National Liberation of East Timor)
Fokupers	Forum Komunikasi Untuk Perempuan Loro Sae (The East Timorese Women's Communication Forum)
Fretilin	Frente Revolucionária do Timor Leste Independente (Revolutionary Front for an Independent East Timor)
GCD	Graveyard Census Database
HRDAG	Human Rights Data Analysis Group
HRVD	Human Rights Violations Database
ICRC	International Committee for Red Cross
ICTJ	International Center for Transitional Justice
IRR	Inter Rater Reliability
Kodam	Territorial Units of the Armed Forces of the Republic of Indonesia
Kodim	Military District Commands of the Armed Forces of the Republic of Indonesia
Kopassus	Special Forces of the Armed Forces of the Republic of Indonesia
Koramil	Military Sub-district Commands of the Armed Forces of the Republic of Indonesia
Korem	Military Resort Commands of the Armed Forces of the Republic of Indonesia (Each KODAM is divided into Military Resort Commands (KOREM) with one battalion each)
MSE	Multiple Systems Estimation
NGO	Non Governmental Organization
POB	Place of Birth
POD	Place of Death
PPS	Probability Proportional to Size
RMS	Retrospective Mortality Survey
TNI	Tentara Nasional Indonesia (Armed Forces of the Republic of Indonesia), which was formerly known as Angkatan Bersenjata Republik Indonesia (ABRI)
UDT	União Democrática Timorense (Timorese Democratic Union)
UN	United Nations
UNAMET	United Nations Assistance Mission in East Timor
UNTAET	United Nations Transitional Administration in East Timor
WNAR	Western North American Region, International Biometric Society

Appendix 2 – List of Figures

Appendix 3 – List of Tables