**Notes**

I appreciate you taking the time to read my work in progress paper, and I am thankful for any and all comments and critiques. Here are a few areas that I am interested in your thoughts on:

* Literature review
  + Does any portion of my lit review feel too in the weeds or out of place?
  + Are there articles I should be including (especially on human rights data) that are missing?
  + Am I going too in depth on how SVAC prevalence data is coded?
* My research question and hypotheses
  + Can you tell what I am asking/trying to achieve?
  + Are my hypotheses too descriptive?
  + I am currently working on strengthening this portion of my paper and would like to hear any suggestions
* Methods
  + Do my proposed variables make sense?
  + What should I do with these variables once I’ve coded them?
  + I am currently working on strengthening this portion of my paper and would like to hear any suggestions

I often fear that I don’t make sense, but then my worry is that I overexplain. I would love to know where you are confused and where you are bored.

Thank you again for reading.

Explaining Discrepancies in Measures of Wartime Sexual Violence

Esmeralda López

March 2023

**Abstract:** Government agencies, UN agencies, and NGOs are increasingly concerned about wartime sexual violence and require systematic data for policymaking. Observability issues and unreliable data have prevented the development of a reliable cross-national dataset documenting the exact count of sexual violence incidents, but estimations of the prevalence of sexual violence are available in the Sexual Violence in Armed Conflict and the Repertoires of Sexual Violence in Armed Conflict datasets. These datasets use text evidence from three sources – the United States State Department, Amnesty International, and the Human Rights Watch – to make their estimations, but there are possible discrepancies between sources and the “zero observations” category (no reported sexual violence) is large and possibly heterogenous. This paper seeks to improve conflict-related sexual violence data collection by (1) evaluating the exact nature of the decision to code an observation as zero, (2) identifying the different possible rationales behind this decision, and (3) systematically analyzing patterns in the zero codings and source discrepancies. I expect to find that the sources’ information is more often nonspecific than nonexistent and that the sources rarely describe the same event, though their descriptions of sexual violence may still overlap broadly.

**Keywords:** conflict-related sexual violence, SVAC, RSVAC, human rights measurement

**Introduction**

What we know about conflict-related sexual violence has increased drastically in the past decade. Amongst these advances in scholarship include the creation of the Sexual Violence in Armed Conflict dataset (Cohen and Nordås, 2014) and the publication of the large number of studies using this data (Nordås and Cohen, 2021). Government agencies, United Nations agencies, and nongovernmental organizations express growing concern for sexual violence in conflict and request systematic data to predict and prevent wartime sexual violence. These policy concerns necessitate the collection of reliable data across time, place, and conflict actor. While scholars in this field are tasked with strengthening systematic data, observability issues make gathering reliable data on sexual violence inherently difficult.

The Sexual Violence in Armed Conflict dataset, SVAC, presents the most established and comprehensive data on wartime sexual violence to date. SVAC and a related dataset, the Repertoires of Sexual Violence in Armed Conflict dataset, RSVAC (Dumaine *et al.*, 2022), include for each observation (conflict, actor, year) measures of reported prevalence of sexual violence sourced from annual and special reports by the United States State Department, Amnesty International, and the Human Rights Watch.

While the prevalence scores associated with a particular observation are often in agreement with each other, there are a number of cases where the prevalence scores, ranging from zero (no reported sexual violence) to three (“massive” levels of sexual violence), exhibit large- and small-scale discrepancies. In an instance of large-scale discrepancy, one source may depict “massive” levels of sexual violence (coded as a three) for a particular observation while the other two seem to not mention sexual violence at all (coded as zeros) for the same observation. These large discrepancies are made salient by the frequency of zeros in the dataset – 80% of the SVAC 2.0 observations, with data spanning from 1989 to 2015, are zeros.

I am interested in two related concerns in the state-of-the-art SVAC and RSVAC datasets: (1) the discrepancies between sources’ prevalence scores and (2) the large and possibly heterogeneous “zero observations” category. I have identified 58 observations where discrepancies are particularly large (between zeros and threes) for initial examination. Continuations of this project may include analysis of the 98 observations with discrepancies between zeros and twos and the 569 observations with discrepancies between zeros and ones.

To investigate the apparent disagreements regarding the prevalence of these 58 large-scale observations, I will (1) evaluate the nature of the decision to code an observation as zero, (2) identify the rationale behind this decision, and (3) systematically analyze patterns in the zero codings and source discrepancies.

The next three sections of this paper will each respectively review the wartime sexual violence literature, human rights measurement literature, and the SVAC and RSVAC prevalence data. In section four, I will give my hypotheses, and in section five, I will introduce a coding mechanism to address the discrepancy and heterogeneity concerns. Finally, I will conclude with plans for the continuation of this project.

**Conflict-related Sexual Violence**

Over the past twenty years, political scientists have begun to pay more attention to conflict-related sexual violence, a topic that was initially of interest primarily to feminist scholars. The intuition of the early political scientists engaged in this study was that sexual violence in conflict is omnipresent, but more recent scholarship has refuted this claim (Nordås and Cohen, 2021). Even conflicts with high levels of political violence may still have relatively low levels of conflict-related sexual violence (Wood, 2006). Noting this variation and that sexual violence prevalence and practices vary across conflicts, more recent scholarship has emphasized the importance of studying the myriad of circumstances associated with different instances of wartime sexual violence (Cohen, 2013; Cohen and Nordås, 2014; Wood, 2018).

Arguments by Cohen (2013) and Wood (2018) demonstrate the importance of studying variation in patterns of conflict-related sexual violence. Cohen (2013) examines the recruitment patterns of conflict actors accused of sexual violence and finds that greater levels of wartime sexual violence, specifically multi-perpetrator rape, occurs in contexts where armed groups have been forcibly recruited, suggesting that such violence may be a “vehicle” for establishing cohesion amongst the armed actors. Wood (2018) presents another analysis of the relationship between armed group socialization and the occurrence of sexual violence, arguing that certain vertical and horizontal socialization patterns—relating to group cohesion and commander control over armed actors, respectively—contribute to increased severity of sexual violence in conflict. The horizontal socialization findings corroborate the group cohesion analysis presented in Cohen (2013). The vertical socialization findings note that sexual violence is more rampant in contexts where armed actors are unpaid, unsupervised, far from home, or deployed for long periods of time (Wood, 2018). Both analyses by Cohen (2013) and Wood (2018) push back against the depiction by earlier scholarship of wartime sexual violence as certain and unavoidable (e.g., Farwell, 2004; Gottschall, 2004).

The identification of varying patterns and risk factors associated with conflict-related sexual violence, such as that by Cohen (2013) and Wood (2018), help policymakers and activists predict and prevent sexual violence, better address the needs of survivors, and allocate resources to conflicts with sharper risk for sexual violence as a practice of war. To mitigate risk of increased sexual violence in populations displaced by conflict, for example, refugee camps should be strategically organized and essential goods such as water, food, and fuel should be easily accessible (Marsh, Purdin and Navani, 2006). The close documentation of wartime sexual violence is also crucial in the rare event that such crimes are prosecuted at international criminal tribunal (de Brouwer, 2015).

SVAC was first introduced in 2014 to fill the need for disaggregated data in the study of conflict-related sexual violence. With 10,954 observations from 1989 to 2015, SVAC documents the conflict and actor type, region, prevalence described by each of three sources, and form of violence for each conflict actor in each inter- and intrastate conflict during each active and interim conflict year as well as for five years post-conflict (Cohen and Nordås, 2014).

RSVAC was released just last year to supplement the SVAC dataset. While the observation unit in SVAC is conflict-actor-year, RSVAC elaborates on the form of sexual violence by organizing each of the eight forms (rape, sexual slavery and forced marriage, forced prostitution, forced pregnancy, forced abortion and forced sterilization, sexual mutilation, sexual abuse, and non-penetrative sexual torture) into separate detailed worksheets of the RSVAC workbook. Each worksheet contains qualitative notes including information on incidents by multiple perpetrators, victims, targeting, timing, and witnesses (Dumaine *et al.*, 2022). The RSVAC dataset is not used for this project because the additional details on the form of sexual violence fall outside its scope.

**Human Rights Measurement and Data Reliability**

Shame, fear of retaliation, trouble reporting, and other difficulties are known to hinder the accurate collection and measurement of sexual violence data. These difficulties are only exacerbated by conflict conditions (Palermo and Peterman, 2011; Wolitzky-Taylor *et al.*, 2011; Krüger and Nordås, 2020; Nordås and Cohen, 2021). Palermo and Peterman (2011) critique the collection of wartime sexual violence data and cite “macro-level (or institutional)” and “micro-level (or individual)” constraints as the main sources of bias and limitation. Macro-level constraints include logistical barriers and a lack of rigorous research that has resulted from sexual violence research being long “sidelined as a ‘feminist’ research agenda.” Micro-level constraints include the effects that stigmatization and fear have on underreporting. The belief that victimhood or survivorship is a path to additional resources is a micro-level constraint that may lead to possible overreporting (Palermo and Peterman, 2011). Because conflict-related sexual violence tends to be obscured, Krüger and Nordås (2020) argues in favor of latent variable modeling to evaluate the prevalence of wartime sexual violence.

Turning to the collection and measurement of human rights broadly, changing standards of accountability and varying levels of government transparency have been known to complicate assessment of the prevalence of human rights violations (Fariss, 2014; Eck and Fariss, 2018). Beyond the establishment of norms and state cooperation, different information providers—newspapers, human rights organizations, and interviews—also shape the story of state violence that is told. Davenport and Ball (2002) describe this as “distinct ‘views to a kill.’” Notably, human rights organizations are more likely to report on human rights abuses that have a large number of killings (Davenport and Ball, 2002). Organizations documenting human rights practices—including those cited by SVAC: the United States State Department, Amnesty International, and the Human Rights Watch—have faced critique of allegedly biased data collection, though the impact of such bias is contested (e.g., Poe, Carey and Vazquez, 2001). Amnesty International and the Human Rights Watch in particular have been accused of problematically framing sexual violence against men as “torture” rather than sexual violence, a label that the organizations may be more willing to apply to instances of violence against women (Charman, 2018). Fariss (2014) also touches on the framing of human rights issues, noting that some researchers introduce indicators of repression like the Political Terror Scale as “measures of *abuse* instead of *reported abuse*,” an important distinction.

**Overview of the SVAC Prevalence Data**

The SVAC codebook, updated in February 2021, states in bolded and underlined text that “[the] SVAC dataset cannot be used as a means to estimate the number of victims” (Cohen, Nordås and Nagel, 2021). This is due to the difficulty in procuring exact counts of victims or survivors, perpetrators, or even incidents of conflict-related sexual violence (Palermo and Peterman, 2011; Krüger and Nordås, 2020; Nordås and Cohen, 2021). In the absence of reliable exact counts, the SVAC dataset bases its assessment of prevalence of wartime sexual violence on qualitative description and relies on count estimations only if qualitative description is not available. The language associated with each prevalence score as well as the necessary criteria for information from the source to be coded at all is summarized below.

A prevalence score of three (“Massive”) means that the sexual violence was “likely related to the conflict” and “described as ‘systematic’ or ‘massive’ or ‘innumerable.’” The conflict actor may have committed sexual violence as a “means of intimidation,” instrument of control and punishment,” “weapon,” “tactic to terrorize the population,” “terror tactic,” “tool of war,” or on a “massive scale.” If the enumerated language is not used, a count of 1,000 or more reports of sexual violence would trigger an SVAC coder to categorize the prevalence for that source as a three (Cohen, Nordås and Nagel, 2021).

A prevalence score of two (“Numerous”) means that the sexual violence described in the source is still likely conflict-related, but the source does not use as strong language as that required for a coding of three. The source instead described the conflict-related violence with words like “widespread,” “common,” “extensive,” “frequent,” “persistent,” or “recurring.” Its occurrence may have been “common,” “often,” or “in large numbers.” Without using these or similar qualitative descriptors, a source could be assigned a two if it mentions a count of 25 to 999 reports (Cohen, Nordås and Nagel, 2021).

A prevalence score of one (“Isolated”) means that the described sexual violence is still likely conflict-related, but the language used by the source is not as strong as that required for a coding of two or three. The source instead mentions “reports,” “isolated reports,” or that “there continued to be reports” of sexual violence. Without these or similar trigger words, a source that mentions a count of less than 25 reports of sexual violence is assigned a one (Cohen, Nordås and Nagel, 2021).

The SVAC codebook explains that for a source to be assigned a prevalence score of zero (“None”), a report was issued, but the report contained “no mention of rape or other sexual violence related to the conflict” (Cohen, Nordås and Nagel, 2021). However, there are circumstances where sources can be assigned a prevalence score of zero even if they release a report that mentions rape or other forms of conflict-related sexual violence. This happens when the information presented in the report falls outside the limits of the dataset or is otherwise too vague to be accurately coded (Krüger and Nordås, 2020). For example, the sexual violence description may not specify the year, conflict, or associated conflict actor, which would make it impossible for a coder to include the information in the dataset, which requires the specification of those categories.

**Hypotheses**

A prevalence score of one, two, or three is associated with the severity of the source’s language—or, absent qualitative description, the number of reported incidents of sexual violence—but zeros can tell two different stories. A zero could mean that the report did not mention sexual violence (nonexistent data). Alternatively, a zero could mean that information on sexual violence was present but lacked sufficient detail to be included in SVAC, which requires conflict-actor-year for each observation (nonspecific data). Each of the 58 observations that I have identified as exhibiting a large-scale discrepancy is associated with three prevalence scores: one score for each of the three sources (the State Department, Amnesty International, and the Human Rights Watch). This makes for a total of 174 prevalence scores, of which 106, or roughly 61%, are zeros.

Because of the multiple avenues that could lead to a prevalence score of zeros and because the zeros significantly outnumber the threes, I find it necessary to analyze the zeros most closely when examining the 58 discrepancies. I expect that such an analysis will show that most of the zeros are the result of nonspecific rather than nonexistent data. I also expect that most of the lack of specificity will involve vagueness surrounding conflict actors, especially those in rebel groups. For example, the Human Rights Watch may describe a rebel group accused of forced sterilization but fail to identify the specific rebel group. Because the specific actor alleged of abuse is not named, SVAC coders would be unable to include the forced sterilization incident described by the Human Rights Watch in the dataset.

**Nonspecific vs. Nonexistent Hypothesis:** Most zeros are from nonspecific information presented by the source rather than an absence of information presented by the source.

**Armed Actor Specificity Hypothesis:** Most specificity issues will involve conflict actors, especially rebel actors.

SVAC coders are not the only ones associated with the distribution of conflict-related sexual violence data. Before any information is considered for the SVAC dataset, the information must be presented in a report by the State Department, Amnesty International, or the Human Rights Watch. While I do not believe that any of these sources are biased to a degree that would significantly impact their discussion of wartime sexual violence, I do anticipate that these sources will not describe what appear to be the same incidents, though the sources may report on the same form of sexual violence. For example, the State Department and Amnesty International may both report on rape (broad overlap), but the reports will mention different regions (no specific overlap). From this information, SVAC may extrapolate the same conclusion that rape is a massive problem associated with a particular conflict-actor-year, but ultimately, the sources do not present the same information to motivate such a conclusion.

In this way, broad and specific disagreements between the sources are not perceptible through the SVAC dataset, but I expect that if the sources are examined directly, there will be more broad overlap between the sources than specific overlap.

**Source Overlap Hypothesis:** The sources will more often describe the same forms of sexual violence (broad overlap) than they will describe the (seemingly) same incidents (specific overlap).

**Methods**

To test my three hypotheses, I have created three sets of variables. I have been given access to the SVAC conflict manuscripts (CMs) and will be relying on these documents to code the variables that I have created. The CMs include notes from the SVAC coders on their coding decisions for each prevalence variable for each observation and many include the relevant portions of the sources’ original reports.

The first set of variables (see Figure 1) addresses my first hypothesis by distinguishing prevalence scores of zero associated with nonspecific information on sexual violence from those associated with nonexistent information on sexual violence. Because there are three prevalence variables associated with each observation in the SVAC dataset, I have created three corresponding binomial variables: state\_z, ai\_z, and hrw\_z. If the SVAC coders assign a prevalence score of zero and note that the report from the source does not include information on sexual violence, I will code my corresponding variable as zero. If the SVAC coders assign a prevalence score of zero and note that the report mentions insufficient information relating to sexual violence, I will code my corresponding variable as 1. If the SVAC coders did not assign a prevalence score of zero, I will code my corresponding variable as NA.

The second set of variables (see Figure 2) addresses my second hypothesis by recording why excluded information on sexual violence could not be included in the SVAC dataset. There are again three variables that correspond to the three prevalence variables in the SVAC dataset: state\_z\_why, ai\_z\_why, and hrw\_z\_why. This set of variables will be coded as a value from 1 through 7 that represents the reason the coders could not include the data. For example, if the coders’ notes in the CM mention an inability to code the data because the source does not specifically identify the conflict associated with the described sexual violence, I will code my corresponding variable as two. In cases where the SVAC prevalence score is not zero and in cases where the SVAC prevalence score is zero due to an absence of information, I will code my corresponding variable as NA.

The third set of variables (see Figure 3) tracks overlap between the sexual violence information presented by the sources. Because there are three sources considered in the SVAC dataset, I have created three variables to correspond with each relationship between the sources: state\_ai\_overlap, state\_hrw\_overlap, and ai\_hrw\_overlap. Where the previous sets of variables that I have described focus on the SVAC coders’ notes, this set of variables focuses on the source material directly to track the broad and specific overlap between sources. If there is no overlap between two sources, I will code the corresponding variable as 0. If there is some broad overlap—meaning both sources describe the same form of violence, such as forced prostitution, but not the same events—I will code the variable as 1. If there is some specific overlap—meaning both sources seem to describe the same events as indicated by reasonably overlapping details—I will code the variable as 2.

While I am still in the process of collecting data for this project, I intend to utilize a latent variable model to evaluate the latent traits associated with the data. After I have created a model, I intend to identify cases that strongly epitomize or deviate from my theory. These notable cases will allow me to elaborate on my model and its real-world applicability. Finally, I would like to continue this research in conversation with individuals affiliated with the State Department, Amnesty International, and the Human Rights Watch.



Figure 1



Figure 2



Figure 3

**Conclusion**

Evaluating the 58 cases that I have identified as large-scale discrepancies is the first step in a series of endeavors to evaluate the SVAC dataset systematically so that future policymaking and scholarship may more closely analyze conflict-related sexual violence. I have identified limitations of my research and outlined potential next steps below.

My current analysis of the SVAC dataset is limited in four primary ways. First, the variables that I have created are applied only to where there exists a gap between the zero and three prevalence scores. Future research may build on this by expanding the variables I have created to the rest of the dataset and testing whether the same patterns exist. Second, I chose to limit my analysis to only cases where all three sources released reports. Future research may consider where a gap between the zero and three prevalence scores exists, but only two sources have released reports. Third, my analysis of the Source Overlap Hypothesis relies on the CMs rather than the original source material. This is for two reasons: first, the CMs appear to be comprehensive, and second, reevaluating every annual and special report by each of the three sources for each of the 58 discrepancies requires more resources than currently feasible. While I believe that the conclusions to be drawn from the CMs are the same conclusions to be drawn from the original source material, this research is still limited to the conclusions drawn from the CMs. Future research may evaluate the original source material to evaluate the argument I posed in the Source Overlap Hypothesis. Lastly, I limited my analysis to the SVAC dataset because my theory has little to do with the form of sexual violence. Future research may expand this study to include analysis of the RSVAC dataset.

Future scholarship in this area may also choose to breech new territory. Where I focused on the widest gaps, those between zeros and threes, future research could target other prevalence score disagreements. How significant is the distinction between a prevalence of one and a prevalence of two? Future research may also choose to evaluate supposed agreement. Are all prevalence of three the same?

The number of studies published since 2014 that utilize SVAC have allowed the progression of wartime sexual violence research. By continuing to analyze and scrutinize existing data on conflict-related sexual violence, we can continue to push the field forward. Efforts to prevent and address sexual violence in conflict through policymaking, physical security measures, strategic allocation of essential goods, and prosecution may also prove more successful with improved systematic data.

**References**

Ballesteros, A. *et al.* (2007) *The Work of Amnesty International and Human Rights Watch: Evidence from Colombia*.

de Brouwer, A.-M. (2015) ‘The Importance of Understanding Sexual Violence in Conflict for Investigation and Prosecution Purposes’, *Cornell International Law Journal*, 48(3), pp. 639–666.

Charman, T. (2018) : ‘The framing of sexual violence against men in armed conflict in Amnesty International and Human Rights Watch reports’, in *Sexual Violence Against Men in Global Politics*. Routledge.

Cohen, D.K. (2013) ‘Explaining Rape during Civil War: Cross-National Evidence (1980–2009)’, *American Political Science Review*, 107(3), pp. 461–477. Available at: https://doi.org/10.1017/S0003055413000221.

Cohen, D.K. and Nordås, R. (2014) ‘Sexual violence in armed conflict: Introducing the SVAC dataset, 1989–2009’, *Journal of Peace Research*, 51(3), pp. 418–428.

Cohen, D.K., Nordås, R. and Nagel, R.U. (2021) ‘Sexual Violence in Armed Conflict Data Project (SVAC) 3.0, 1989-2019 Codebook and Instruction Manual February 2021’, p. 17.

Davenport, C. and Ball, P. (2002) ‘Views to a Kill: Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977-1995’, *The Journal of Conflict Resolution*, 46(3), pp. 427–450.

Dumaine, L. *et al.* (2022) ‘Repertoires of conflict-related sexual violence: Introducing the RSVAC data package’, *Journal of Peace Research*, 59(4), pp. 611–621. Available at: https://doi.org/10.1177/00223433211044674.

Eck, K. and Fariss, C.J. (2018) ‘III Treatment and Torture in Sweden: A Critique of Cross-Case Comparisons’, *Human Rights Quarterly*, 40(3), pp. 591–604. Available at: https://doi.org/10.1353/hrq.2018.0033.

Fariss, C.J. (2014) ‘Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability’, *American Political Science Review*, 108(2), pp. 297–318. Available at: https://doi.org/10.1017/S0003055414000070.

Farwell, N. (2004) ‘War Rape: New Conceptualizations and Responses’, *Affilia*, 19(4), pp. 353–474. Available at: https://doi.org/10.1177/0886109904268868.

Gottschall, J. (2004) ‘Explaining Wartime Rape’, *The Journal of Sex Research*, 41(2), pp. 129–136.

Krüger, J. and Nordås, R. (2020) ‘A latent variable approach to measuring wartime sexual violence’, *Journal of Peace Research*, 57(6), pp. 728–739. Available at: https://doi.org/10.1177/0022343320961147.

Marsh, M., Purdin, S. and Navani, S. (2006) ‘Addressing sexual violence in humanitarian emergencies’, *Global Public Health*, 1(2), pp. 133–146. Available at: https://doi.org/10.1080/17441690600652787.

Nordås, R. and Cohen, D.K. (2021) ‘Conflict-Related Sexual Violence’, *Annual Review of Political Science*, 24(1), pp. 193–211. Available at: https://doi.org/10.1146/annurev-polisci-041719-102620.

Palermo, T. and Peterman, A. (2011) ‘Undercounting, overcounting and the longevity of flawed estimates: statistics on sexual violence in conflict’, *Bulletin of the World Health Organization*, 89, pp. 924–925. Available at: https://doi.org/10.1590/S0042-96862011001200017.

Poe, S.C., Carey, S.C. and Vazquez, T.C. (2001) ‘How Are These Pictures Different? A Quantitative Comparison of the US State Department and Amnesty International Human Rights Reports, 1976-1995’, *Human Rights Quarterly*, 23(3), pp. 650–677.

Wolitzky-Taylor, K.B. *et al.* (2011) ‘Is Reporting of Rape on the Rise? A Comparison of Women With Reported Versus Unreported Rape Experiences in the National Women’s Study-Replication’, *Journal of Interpersonal Violence*, 26(4), pp. 807–832. Available at: https://doi.org/10.1177/0886260510365869.

Wood, E.J. (2006) ‘Variation in Sexual Violence during War’, *Politics & Society*, 34(3), pp. 307–342. Available at: https://doi.org/10.1177/0032329206290426.

Wood, E.J. (2018) ‘Rape as a Practice of War: Toward a Typology of Political Violence’, *Politics & Society*, 46(4), pp. 513–537. Available at: https://doi.org/10.1177/0032329218773710.