Instructions to replicate the analyses in the manuscript and Appendix of "Who, what, and where? Linking violence to civil wars"

Software required for replication: R

Code for the replication is: analysis_replication_code.R

Necessary datasets for the replication are all provided in the ZIP folder "R&P Analysis Replication". The user needs to download them on a local repository or server and set the working directory in R to where they have stored those. Here is a list and brief description of these datasets:

- **GED-to-ACD_conflict_month_final_v241.rds**: Our own data described in the article, created with our functions and all default choices, except for including all postwar months.
- **div_GED-to-ACD_conflict_month_final_v241.rds:** The same dataset but created with user choice (in the data creation) to divide deaths between conflict zones in the case of zone overlap.
- **unclipped_GED-to-ACD_conflict_month_final_v241.rds:** The same dataset, but created with user choice (in the data creation) to not clip the conflict zones at the country border.
- **PRIO100_replication_final.dta**: Data used in a study by Balcells and Kalyvas (2014), with information on type of warfare (conventional, irregular, and symmetric nonconventional). Downloaded from: https://www.laiabalcells.com/pages/5469
- **RAD_grouplevel_data.csv**: Data by Mehltretter et al (2023) on rebel access to Major Conventional Weapons (MCW). Downloaded from: https://www.gsi.uni-muenchen.de/lehreinheiten/le_policy_analysis/forschung/armstrade_intrastate/index.html
- **translate_conf.csv**: Conversion table provided by the Uppsala Conflict Data Program to translate between older conflict IDs used for instance by Balcells and Kalyvas (2014), and the newer ones, used for instance in our analysis. Downloaded from https://ucdp.uu.se/downloads/
- **ged_prepared_totracklinks_v241.rds**: UCDP GED with some data pre-processing steps as indicated in our data creation code, i.e., multi-month events divided into separate months, violence by coalitions divided into the individual perpetrators, etc. In this file, our code documents how each event is attributed to a conflict (e.g., via actor, via space).
- **Intersection files**: b0_gedintersection_v24.rds, b10_gedintersection_v24.rds, b20_gedintersection_v24.rds, b30_gedintersection_v24.rds. Result of an in-between step in our data creation code (the moment GED is intersected with the conflict zones), with different buffer user choices. Used for sensitivity tests in the Appendix.
- **Conflict zones**: Shapefiles of the conflict zones in 5 versions, 4 with different buffers, and one final with conflict zones clipped at the country border.

Once the necessary packages are installed and loaded (code provided), the user can produce the respective figures and graphs by just following the code from top to bottom.

Comments on Figure 4b

The figures were created by turning our final data into annual data, and then merging info on type of warfare from Balcells and Kalyvas (2014) and on rebel access to Major Conventional Weapons from Pamp (2024) into our data. This allows us to assess the gap between battle deaths and overall intensity for different types of conflicts. We chose to use both datasets as the Balcells and Kalyvas data ends in 2008 and only includes UCDP conflicts that caused over 100 deaths a year. The graph on the left accordingly captures only a small subset of conflict-years in our data. We thus also use the data by Pamp et al. with coverage until 2020.

References

Balcells, Laia, and Stathis N. Kalyvas. "Does warfare matter? Severity, duration, and outcomes of civil wars." *Journal of Conflict Resolution* 58.8 (2014): 1390-1418.

Pamp, O., Mehltretter, A., Binder, P., & Thurner, P. W. (2024). Introducing the Rebels' Armament Dataset (RAD): Empirical Evidence on Rebel Military Capabilities. *Journal of Conflict Resolution,* online first.