**CHAPTER THREE**

**Dataset and Quantitative Methods**

*Section 1: Dataset Introduction and Overview*

IThere were multiple non-trivial issues regarding the study of human rights violations: For one, prior to this study there existed no dataset of human rights violations with respect to civil wars, much less civil war duration. In addition, data of human rights violations can often be incomplete, vague, or imprecise. As a result, to conduct this study I have primarily had to incorporate data from three datasets to create a new dataset and properly evaluate the relationship between human rights violations and civil war duration: The Political Instability Task Force (PITF) Political Atrocities dataset, the Correlates of War (COW) Intrastate-War dataset, and Fearon and Laitin’s ‘Ethnicity, Insurgency, and Civil War’ replication dataset. In addition, Fearon and Laitin’s dataset was truncated at the year 2000; as such, I had to refer to the World Bank and the United Nations for datasets on year-indexed GDP per capita and population, respectively, for all countries included in my dataset. Some data cleaning and investigation was required to account for missing and vague data; all methods and statistical coding used to combine and clean the included datasets can be found in Appendix A.

In this section, I will begin by discussing the curation of civil wars included in this study, accounting for why each conflict was chosen. This is followed by an explication of the primary variables investigated in each of the three incorporated datasets and justifying their inclusion in this study, as well as account for custom variables that I developed. Subsequently, I will examine potential shortcomings of the dataset, including potential sources of bias, and account for any adjustments or corrections I’ve made for the purposes of this analysis. To conclude, I will qualify my use of Hypothesis Testing and Cox Regression analysis as a means to complete survival analyses of the incorporated explanatory variables.

*Section II: Civil War and Human Rights Violation Curation*

It appears we can define civil wars broadly as *wars that occur within the political boundaries of a state*. However, academic discussion on the definition of civil wars – and their subsequent inclusion into conflict studies – has not been so certain. As Testerman writes in his analysis of rebel financing during civil wars, “news of violent internal [state] conflict frequently turns to the question, is this a civil war?”[[1]](#footnote-1). Nicholas Sambanis concurs that the question is complex and finds that varying definitions of civil war can lead counts of disparate intrastate conflicts to range from 58 to 116, depending on the dataset used[[2]](#footnote-2). Of the dozen datasets included in Sambanis’ analysis, correlation coefficients spanned values from .42 to .96 (Sambanis 832 – 834), suggesting a lack of consensus regarding the definition of a civil war[[3]](#footnote-3). If the curation of wars into a dataset is not qualified, any related analyses could be misleading by containing diverse combinations of conflicts.

1. Testerman, Matthew A. “Rebel Financing in Civil Wars: A Quantitative Analysis with Three Case Studies of Civil Wars in the Phillippines.” 47. 2012. [↑](#footnote-ref-1)
2. Sambanis, Nicholas. “What is Civil War?” 832-834. 2004. [↑](#footnote-ref-2)
3. See footnote above. [↑](#footnote-ref-3)