Measuring Sub-national Levels of Repression: A Latent Variable Model[*]

Rebecca Cordell[†]

K. Chad Clay[‡]

Christopher J. Fariss[§]

Reed M. Wood[¶]

Thorin M. Wright[‖]

Abstract

We present a measurement model that aggregates human rights information from country-year allegation data taken directly from the human rights reports published annually from Amnesty International, Human Rights Watch, and the US State Department. First, we review categorized information about four specific types of physical integrity violations (disappearances, torture, killing and political imprisonment). In addition each allegation is coded for content that includes the date, actor type, location, scope, intensity, range, victim type, and whether or not the violator of the right was held accountable. This allegation dataset is generated from a supervised machine learning coding process and contains over 100,000 human rights allegation observations. Second, we develop latent variable models that aggregate this information. The measurement models are hierarchical because each version aggregates information from these allegation records into estimates that allow for the comparison of country-year and sub-national-year units. Because we have disaggregated the content of the annual human rights country reports, the measurement models are able to use any single piece of information or combination of information into estimates that are values for higher level country and sub-country units. We compare these new country-year latent variable estimates to existing ones and the new sub-country-year units to other disaggregated information related to human rights abuse.