We had a good conversation today, but there’s a challenge we need to solve before we do a ton of work on mapping out how to code events (or anything else). Our conception of the target variables (DV’s) was that we would focus on three behaviors by the state that represented democratic backsliding or closure of civic spaces:

1. Violence / incarceration of journalists or other media / political actors.
2. Limitations on speech – e.g., institutional changes that hampered the ability of the press to investigate and report, turning off the internet, …
3. Limitations on association – e.g., implementing curfews, laws against assembly, …

We’ve now looked at several cases and the main challenge is that most of these measures are very sparse. E.g., in an entire year in Kenya (which features real drama wrt elections and the presidency), there were five (5) cases of physical harm or harassment of journalists. And for institutional changes, once they happen (e.g., Russia’s semi-recent law to prevent foreign ownership of media) they aren’t by definition going to happen again. Bottom line: we do not have high frequency data on anything we care about for the DV’s.

Potential options:

1. We could commit to studying this in a comparative frame – i.e., all nations of interest will be lumped together in one sample. A DV, e.g., on jailings / harassment of journalists might be low in most cases but if it’s high in a subset, that might well be informative about future backsliding / closures. Downside: we’d have to arrive at a set of measures that are general across cases and believe that one model fits all.
2. We could form an index of “bad behaviors” by the government, including internet shutdowns, jailings, institutional changes, etc. This would provide enough variance in the DV to build models. Downside: I don’t see a non-vomitous way to do this given the types of events, the fact that they aren’t all equally frequent, and that they aren’t all of the same magnitude in terms of effect on civic spaces.

We need to solve this issue, imo, before we do anything else. There’s no chance of building a good forecasting model without high quality target variables.