Hi Simon,

This is much better, you've evidently worked hard on it. Here are some comments:

Paper 1:

- You use Druckman and Chong as an example of small treatment groups and thus risks to balance (p.1). But you suggest it without showing it. Are there important imbalances here or aren't there? If you can't demonstrate with data, don't use it as an example. Find something else, run a simple simulation, whatever.

- You claim that there is no method to sequentially block on ordinal measures (p.2). Why is that? What is the difficulty, specifically?

- You make the small groups risks no balance claim again on p.4. Do we have any evidence that this happens *in public opinion studies*? You need to make this case because the theme of your dissertation is, after all, improving survey research. You still haven't shown me that this theoretical risk has undermined a single actual study of public opinion.

- You don't actually explain what blocking does in any detail (p.5). Beyond, that is, simply stating that we are somehow "using covariate information", which is not terribly helpful because that's more or less what regression and all other multivariate analysis does.

- Some real fuzziness here in your discussion of blocking v. OLS controls (p.5). There's a) no real specificity about what the actual algorithms are (see prev. point); b) there is a distinction between the ends (balance) and the means (regression assumptions v. blocking) that is unexplained and unexplored; c) the point about controls should in theory apply to both approaches, e.g. either way you need to know what the confounds are; d) "blocking balances variation…" sentence is not explained in any useful way.

- In the sequential blocking bit (p.6), you start by outlining two scenarios, the second of which being the one suited to sequential blocking. Show me what survey experiments, ideally in political science but even more generally, fall under this category. This matters because, as noted above, your justification for the third paper and indeed the whole dissertation is that it applies the methods in a way that fixes survey research.

- Your simulation (p.16) is described in only the vaguest possible terms. 1) Mean of what? Mean difference between groups? 2) What is the covariate you're matching on, e.g. how many different categories will there be? 3) How many comparisons will there be per simulation "run"? If you have five treatment groups you should have 13 unique mean differences for each run and matching covariate, correct? (all the unique comparisons of groups 1-5). Basically I want to understand this in more detail and this leaves a lot to the imagination.

Paper 2

- The sections describing entropy (p.20, 28) don't really explain the problem with more standard measures of spread such as variaance, skew, kurtosis, etc. The point of developing the entropy measure is that these things are useless w/r/t nominal-level variables, especially if there are more than two categories. But once you get to ordinal-level measurement, you don't *really* have this issue anymore. (Yes, I know that technically you need interval-level measures for variance et al but come on.) Point is, your justification for applying entropy to ordinal+ level stuff strikes me as self-defeating because you're back into territory where standard measures of spread are useful. Please clarify this and provide further justification.

- You still have no substantive expectation of (or justification for) what the effect of the specific mode differences will be on entropy. That is, you neither say what exactly it is about mode that seems to matter in the Homola et al paper or, for that matter, the direction of the effect. Nor do you say anything about what you

expect to happen with ordinal-level measures or why.

Paper 3

- you're still (as ever) out on a ledge in assuming that either method from Papers 1/2 will change your substantive findings in any way.

- your hypothesis is still not properly worded as such (p.38). This is because "frame strength" is still basically a concept here rather than an operationalized variable. Should be: moralized frames cause Y - where Y is whatever your outcomes are) more than non-moralized frames do or something like that.

- I think you've underestimated the difficulty of meta-analysis in this application, as the sample of studies you are likely to pull in are extremely heterogeneous on a number of dimensions. But ok.

- we need to talk more at some point about getting peoples' assessments about what is "moral" and what isn't. My suspicion is that most people will take moral to mean simply something "good" or something "I like" rather than as an attempt to outline a category of argument. Anyway, we can talk about that later.

- I'm unclear about what "Online Poll #2" is, especially given that you say it will yield "qualitative insights". What exactly did you have in mind, here? In the last ground you talked about focus groups, this seems like you're doing an online survey. If so, very different and tricky in its own way.

Upshot:

Let me know if any of the above is unclear. Ultimately this depends on Jeff, but I'm willing to sign the advancement paperwork in spite of these shortcomings (e.g. I'm not especially interested in demanding another round of revisions before getting you on your way with this). However, I strongly encourage you to keep them in mind as you progress. These are things that will certainly come back around when the defense comes if not addressed on some level.

Cheers,

M@

...