

POL-GA 1251
Quantitative Political Analysis II
Homework 4

You are going to work with data from the following study:

Gibson, James L. 2004. "Does Truth Lead to Reconciliation? Testing the Causal Assumptions of the South African Truth and Reconciliation Process." *American Journal of Political Science* 48(2):201-217.

These data come from a 2000-2001 survey in South Africa that Gibson used to study inter-racial attitudes ten years after the release of Nelson Mandela and seven years after the elections of 1994 that marked the formal end of the apartheid era. Among other things, he studied the impact of the national Truth and Reconciliation Commission to promote inter-racial reconciliation. The full analysis of the survey appears in his book, *Overcoming Apartheid: Can Truth Reconcile a Divided Nation?* (Russell Sage, 2004). The replication materials are available from the ICPSR website, although they are in SPSS and SAS file formats. So to make things easier, I have converted them to Stata .dta format. I have also provided the codebook from the ICPSR site. For the sake of this exercise, you can just drop any observations where necessary variables have missing, "don't know," "refused," or "no answer" responses (be careful to filter those out).

You are going to use matching to analyze how exposure to the TRC affected attitudes toward apartheid and reconciliation. In R, you could use the package "MatchIt," which provides options for doing many different types of matching, including nearest neighbor matching, GenMatch, CEM, etc. In Stata, you could use one of the matching estimators in the "teffects" command suite. You are free to decide what kind of matching you want to do. Here are the steps:

1. First, construct two outcome variables (2.5 points):

- Construct Gibson's "reconciliation index" as discussed on pages 205-206 of the *AJPS* paper. The relevant variables (as described in the caption to Table 1 in the paper) are as follows:

RUSTAND

RFRIEND2

RCRIME

RTRUST

RSELF

RUNCOMP

RBELIEV

RNONE

RPARTY

(For these variables, Gibson doesn't indicate in the paper whether he counts "uncertain" as a "reconciled" or "non reconciled" answer. Make a decision for yourself and then use that as a way to construct the index.)

- Construct Gibson’s “truth acceptance index” as discussed on pages 205-207 of the paper. The relevant variables, as described in Table 1 are:

TRUTH6

TRUTH1

TRUTH4

ATROC2

ATROC3

In the 2004 book, Gibson clarifies that the index is created by counting “the number of truths about apartheid endorsed by the respondent” (p. 85). It would seem that “uncertain” answers would not count as endorsing the relevant “truth” (see the caption to Table 1 to understand how to code responses as endorsing the “truth.”)

2. Next, construct the treatment (2.5 points):

- Our measure of “exposure to the TRC” will be the variable TRCKNOW, which, admittedly is imperfect as it is a subjective indicator that may be quite error prone. Nonetheless, it is the best available in the data.
- TRCKNOW is not a binary variable. For the sake of this exercise, you can construct a binary variable by distinguishing between those who know either “a great deal” or “some, but not a great deal” (consider them “treated”, and so assign a treatment value of 1) from those who know “not very much” or “nothing at all” (consider them “control”, so assign a treatment value of 0).

3. Use your substantive judgment to establish a set of control variables for your analysis (5 points):

- Explain your selections. What factors might confound the relationship between TRC exposure (as measured) and potential outcomes for either truth acceptance or racial reconciliation? You should obviously include race. But there are others you would probably want to include too. Be thoughtful about this. Keep post-treatment biases in mind when determining what to include and what to exclude.
- You can use the variables in their raw form (excluding the values corresponding to “don’t know,” “refused,” etc., of course), create dummy variables, or create scales (e.g., adding variable values together or using factor analysis, etc.) as you see fit.
- Discuss how plausible is CIA given the control variables that you have available, and what kinds of things you would have liked to measure, if you were to do the survey over again.

4. Construct a matched dataset that controls for the control variables that you selected. Present tables or graphs showing where you made improvements in covariate balance. Discuss this briefly. Based on areas of common support for your covariates, are you proposing to estimate an ATT, ATC, or some other local average treatment effect? (10 points)

5. Estimate the effect of TRC exposure on both Truth Acceptance and Racial Reconciliation.

For now, you can do so by simply working with the matched data and regressing the outcomes on the treatment indicator (equivalent to taking the difference in means), being sure to incorporate any matching weights that your matching algorithm requires and using heteroskedasticity-robust standard errors. Do this for the whole matched sample, and then estimate separate effects for the matched Black/African, White, Coloured, and Indian/Asian subsamples. Present these in publication-quality tables. (In future homework you will do more to probe the accuracy of these standard errors and inferences.) (10 points)