Instructions for Replicating the Simulations

The programs assume the following:

1. You have a subdirectory called STATA on your c:\ directory.
2. The STATA directory has three subdirectories: data, do, and log.
3. Both data and log have subdirectories LK5set.

Thus the following directories are assumed to exist:

c:\STATA\data
c:\STATA\data\LK5set

c:\STATA\do

c:\STATA\log
c:\STATA\log\LK5set

All the "do" files (those with a "do" suffix) should go in the STATA\do directory, while those with a "raw" suffix should go in the STATA\data directory. The programs will write their useful output files in the LK5set directories.

Political Analysis conventions for web based files require a prefix that both identifies the first author and identifies that the files are to be included on the web. In this case the prefix combination is "macdonald.web." This prefix should be removed when the "do" and "raw" files are put on their respective directories.

Obviously you can adopt the directory information to suit your own preferences, but the various do routines are called in the course of the simulations. The names in the program must match the correct files on your computer.

The programs should be easy to run, but they were not designed as mass user programs. Unless you modify the code, you will need to start up STATA each time you run one of the simulations. If you do this, the programs should run and produce the results we report in the article.

If you have any problem running the simulations, please send an email to Rabinowitz (email address: rabinowitz@unc.edu).

The specific instructions for running the simulations are on the next page.

Instructions for Running the Simulations

There are a number of do files in the set of files associated with the simulations, but the only ones you need to use are **acmainsim.do** and **lsmainsim.do**. These are the main programs for the assimilation and contrast simulations and the linear scale simulations, respectively.

The **maxsim** line sets the number of simulations to be run. Currently maxsim is set at 500 because that is the number of each type of simulation we ran for the article. You might want to reset maxsim to a low number (e.g., maxsim = 3) to be sure the program is running well before you run the full set of simulations.

Three different simulations run from the main programs: directional, squared Euclidean proximity, and city-block proximity. Currently the command line to run each of the simulations is preceded by an asterisk. You just need to remove the asterisk from the simulation you want to run, and leave the asterisks on the other two.

```
*do c:\stata\do\acdirsimphaseLK5;     /*directional*/
*do c:\stata\do\acproxCBsimphaseLK5;    /*city-block proximity*/
*do c:\stata\do\acproxEU2simphaseLK5; /*squared Euclidean proximity*/
```

REMEMBER, TO RUN ANY SIMULATION YOU MUST CLOSE AND RESTART STATA!

After a simulation has run, there is a program for summarizing the output. This is the set of do files starting with the prefix "summarize." Simply run the summarize program appropriate for the simulation you have run. So, for example, if you have run the assimilation and contrast directional simulations, run the program **summarizeacdirLK5.do**.

This will summarize the results and save the summary in both SMCL and LOG format on the STATA/log/LK5set directory.