Listed under Homework #1 on our class page is a link to the Gerber and Green article on our reading list, as well as the replication data for this article in comma delimited format. In the replication data all variable names should be self-explanatory, except that `persngrp` and `phonegrp` indicate assignment to the personal contact and phone group treatments, and `persncont` and `phonecont` indicate successful contact with the subject.

(1) Read the Gerber and Green article, and write a paragraph summarizing the field experiment they conducted. In particular, describe the potential endogeneity problem that Gerber and Green discuss, and their proposed solution to this problem.

(2) Using the data provided, replicate the two-stage least squares results presented in Table 5. Do not worry about including the ward dummy variables or clustering on the standard errors. Note: the results will vary slightly since we are using a simplified model, and Gerber and Green corrected some errors in the data since the publication of their article).

(3) Estimate a simple linear regression version of the models in (2), without correcting for endogeneity. How do these results differ from your results in (2). What statistical problems might these results have?

(4) Now replicate your results from (2) “by hand” – that is, by running the first stage regression, generating predicted values, and using these predicted values in a second stage regression. Explain how these results compare to your results in (2).