Election and Conditional Cash Transfer Program in Mexico

5 August 2015

In this exercise, we analyze the data from a study that seeks to estimate the electoral impact of 'Progresa', Mexico's *conditional cash transfer program* (CCT program). This exercise is based on the following article: Ana de la O. (2013). 'Do Conditional Cash Transfers Affect Voting Behavior? Evidence from a Randomized Experiment in Mexico.' *American Journal of Political Science*, 57:1, pp.1-14. and Kosuke Imai, Gary King, and Carlos Velasco. (2015). 'Do Nonpartisan Programmatic Policies Have Partisan Electoral Effects? Evidence from Two Large Scale Randomized Experiments.' Working Paper.

The original study relied on a randomized evaluation of the CCT program in which eligible villages were randomly assigned to receive the program either 21 (Early *Progresa*) or 6 months (Late *Progresa*) before the 2000 Mexican presidential election. The author of the original study hypothesized that the CCT program would mobilize voters, leading to an increase in turnout and support for the incumbent party (PRI in this case). The analysis was based on a sample of precincts that contain at most one participating village in the evaluation.

The data we analyze are available as the CSV file `progresa.csv`. The names and descriptions of variables in the data set are:

Name	Description
`treatment`	Whether an electoral precinct contains a village where households received Early *Progresa*
`pri2000s`	PRI votes in the 2000 election as a share of precinct population above 18
`pri2000v`	Official PRI vote share in the 2000 election
`t2000`	Turnout in the 2000 election as a share of precinct population above 18
`t2000r`	Official turnout in the 2000 election
`pri1994`	Total PRI votes in the 1994 presidential election
`pan1994`	Total PAN votes in the 1994 presidential election
`prd1994`	Total PRD votes in the 1994 presidential election
`pri1994s`	Total PRI votes in the 1994 election as a share of precinct population above 18
`pan1994s`	Total PAN votes in the 1994 election as a share of precinct population above 18
`prd1994s`	Total PRD votes in the 1994 election as a share of precinct population above 18
`pri1994v`	Official PRI vote share in the 1994 election
`pan1994v`	Official PAN vote share in the 1994 election
`prd1994v`	Official PRD vote share in the 1994 election
`t1994`	Turnout in the 1994 election as a share of precinct population above 18
`t1994r`	Official turnout in the 1994 election
`votos1994`	Total votes cast in the 1994 presidential election
`avgpoverty`	Precinct Avg of Village Poverty Index
`pobtot1994`	Total Population in the precinct

Name	Description
`villages`	Number of villages in the precinct

Each observation in the data represents a precinct, and for each precinct the file contains information about its treatment status, the outcomes of interest, socioeconomic indicators, and other precinct characteristics.

Question 1

Estimate the impact of the CCT program on turnout and support for the incumbent party (PRI or Partido Revolucionario Institucional) by comparing the average electoral outcomes in the 'treated' (Early *Progresa*) precincts versus the ones observed in 'control' (Late *Progresa*) precincts. Next, estimate these effects by regressing the outcome variable on the treatment variable. Interpret and compare the estimates under these approaches. Here, following the original analysis, use the turnout and support rates as shares of the voting eligible population (`t2000` and `pri2000s`, respectively). Do the results support the hypothesis? Provide a brief interpretation.

Question 2

In the original analysis, the authors fit a linear regression model that includes, as predictors a set of pre-treatment covariates as well as the treatment variable. Here, we fit a similar model for each outcome that includes the average poverty level in a precinct (`avgpoverty`), the total precinct population in 1994 (`pobtot1994`), the total number of voters who turned out in the previous election (`votos1994`), and the total number of votes cast for each of the three main competing parties in the previous election (`pri1994` for PRI, `pan1994` for Partido Acción Nacional or PAN, and `prd1994` for Partido de la Revolución Democrática or PRD). Use the same outcome variables as in the original analysis that are based on the shares of the voting age population. According to this model, what are the estimated average effects of the program's availability on turnout and support for the incumbent party? Are these results different from what you obtained in the previous question?

Question 3

Next, we consider an alternative, and more natural, model specification. We will use the original outcome variables as in the previous question. However, our model should include the previous election outcome variables measured as shares of the voting age population (as done for the outcome variables `t1994`, `pri1994s`, `pan1994s`, and `prd1994s`) instead of those measured in counts. In addition, we apply the natural logarithm transformation to the precinct population variable when including it as a predictor. As in the original model, our model includes the average poverty index as an additional predictor. Are the results based on these new model specifications different from what we obtained in the previous question? If the results are different, which model fits the data better?

Question 4

We examine the balance of some pre-treatment variables used in the previous analyses. Using boxplots, compare the distributions of the precinct population (on the original scale), average poverty index, previous turnout rate (as a share of the voting age population), and previous PRI support rate (as a share of the voting age population) between the treatment and control groups. Comment on the patterns you observe.

Question 5

We next use the official turnout rate `t2000r` (as a share of the registered voters) as the outcome variable rather than the turnout rate used in the original analysis (as a share of the voting age population). Similarly, we use the official PRI's vote share `pri2000v` (as a share of all votes cast) rather than the PRI's support rate (as a share of the voting age population). Compute the average treatment effect of the CCT program using a linear regression with the average poverty index, the log-transformed precinct population, and the previous official election outcome variables (`t1994r` for the previous turnout; `pri1994v`, `pan1994v`, and `prd1994v` for the previous PRI, PAN, and PRD vote shares). Briefly interpret the results.

Question 6

So far we have focused on estimating the average treatment effects of the CCT program. However, these effects may vary from one precinct to another. One important dimension to consider is poverty. We may hypothesize that since individuals in precincts with higher levels of poverty are more receptive to the cash transfers, they are more likely to turn out in the election and support the incumbent party when receiving the CCT program. Assess this possibility by examining how the average treatment effect of the policy varies by different levels of poverty for precincts. To do so, fit a linear regression with the following predictors: the treatment variable, the log transformed precinct population, the average poverty index and its square, the interaction between the treatment and poverty index, and the interaction between the treatment and squared poverty index. Estimate the average effects for unique observed values and plot them as a function of the average poverty level. Comment on the resulting plot.