Quantitative Political Analysis II, Homework 1

Alessandro Vecchiato

February 5, 2014

Exercise 1

The effect is not identified. This is because the endogenous treatment function is not informative of the treatment status for never-takers and always-takers. Formally this can be proven by noticing that the two estimators for $Z_i = 1$ and $Z_i = 0$ can be represented by:

$$ATE_{Z=1} = E[Y_{i1}|D_i = 1] \underbrace{Pr[D_i(1) = 1]}_{compliers+always-takers} - E[Y_{i0}|D_i = 0] \underbrace{Pr[D_i(1) = 0]}_{never-taker}. \quad (1)$$

Analogously,

$$ATE_{Z=0} = E[Y_{i1}|D_i = 1] \underbrace{Pr[D_i(0) = 1]}_{always-takers} - E[Y_{i0}|D_i = 0] \underbrace{Pr[D_i(0) = 0])}_{compliers+never-takers}. \quad (2)$$

Hence, the "stratified estimator" is computed as a weighted average of:

$$\rho = E[Y_{1i} - Y_{0i}] = \sum_{i=0,1} ATE_{Z=i} Pr[Z_i = i] = ATE_{Z=1} Pr[Z_i = 1] + ATE_{Z=0} Pr[Z_i = 0]. \quad (3)$$

Notice that even though we know the probability of being in each stratum, that is $Pr[Z_i = 1]$ and $Pr[Z_i = 0]$, without knowing the probability that within each stratum individuals belong to each group of compliers, always-takers or never-taker, we cannot reconstruct the effect for each group. Moreover, the conditional independence assumption $D_i \perp (Y_{i1}, Y_{i0}|Z_i)$ is not satisfied within each stratum and, generally, the ATE is not identified.

Exercise 2

Each paper is addressed in turn.

Gamm and Kousser(2013)

Q1: The paper tries to detect the effect of city size, as determined by delegation size, on approval rate of bills in State senate.

Q2: Given the amount of confounding variables induced by the specific content of the bill, ideally the study should deliver random assignment of bills to city delegation to break the bias due to either unobservable characteristics of the delegation or of the bill.

Q3: The authors estimate the effect of city size on bills by controlling for each possible confounding variable related to: (1) Bill content, (2) Author characteristics, (3) Legislative and state content. It is assumed that, conditionally on these controls, the deviations on approval rate are independent to the city size. The statistical model is a multilevel logit model, with coefficients nested at the level of the legislative session. Fixed effect for session, year and state are considered to control for autocorrelation in the data.

Q4: The dataset contains 1736 district bills drawn for large and small legislatures, which are a random sample of the total population of bills discussed by state legislatures. Standard errors and parameters are clustered at the session level.

Lyall, Blair and Imai (2013)

Q1: Wartime victimization and attitudes toward combatants conditional on their identity.

Q2: Random assignment of wartime victimization to a sample of civilian population by different groups. To control for decay in the effect, produce multiple victimization assignments to people and comparison of the effect over time.

Q3: The authors conduct a survey of multiple endorsement experiment to measure attitudes. The treatment is elicited opinion toward a policy endorsed by either the ISAF or the Talibans. The control group is elicited with opinion on the same policies without the endorsement of the group. To avoid confounding by the specific policy content, four policies are considered that belong to the same policy space, are widely publicized, were actually endorsed by the groups and widely debated by the general public. The statistical model they use is a ordered probit estimation where the latent parameters are the initial support for the reform and for the group and the dependent variable is the support conditional on different endorsement. The model includes demografic, local and timing control covariates and incorporates cases of respondent's direct and indirect exposure to harm.

Q4: The survey experiment was conducted between 18 January and 3 February 2011 in five provinces of Afghanistan. Sampling is multi-stage so to generate estimates at the individual, village and district levels. The reference population is the whole population of Afganistan. They sample only small rural areas, as the majority of the population in Afghanistan lives in these contexts.

Paler (2013)

Q1: Causal effect of differential ways to finance government budget and government accountability.

Q2: Random assignment of different financing sources to otherwise homogeneous treatment group of state governments.

Q3: The research design includes a revenue experiment of home budgeting in the region of Blora. The treatment was a forced permanent transfer of resources from the household budget board to the government board.In the windfall groups resources were introduced as coming unspecified by other government sources. Then the treatment effect is detected by t-test of difference in means in attitudinal variables like willingness to monitor, political participation and incumbent sanctioning.

Q4: Participants in the campaign were randomly sampled using multistage cluster sampling from the adult population in Blora. The target population was all individuals between the ages of 17 (the voter eligibility age) and 65 who had resided in Blora for at least six months. First, 93 of 295 villages were randomly sampled within strata formed by subdistrict and urban-rural status. Within each village, one subvillage unit (dusun) was randomly sampled, followed by the random selection of 20 households from an updated list of all households in the dusun. Canvassers then sampled one participant in each household on arrival using simple random sampling from a full list of eligible household members made in consultation with a head of household. Random assignment was blocked at the village level so that, of the 20 participants per village, five were randomly assigned to each of the four experimental conditions

Meredith (2013)

Q1: Estimate the extent of cottail effect in gubernatorial, secretary of state and attorney general elections.

Q2: Ideally you would want to randomly assign candidate personal support in each county, and evaluate the effect to down-ballot vote share.

Q3: They use an IV estimation, using as instrument the extent of friends-and-neighbour vote measured by geographical proximity of the ballot to place of birth and residence of gubernatorial, secretary of state, and attorney general candidates. In the second stage they use this exogenous source of variation to estimate the extent of cottail effect on concurring elections.

Q4: The exclusion restriction in this context is that $E[\epsilon_{c,t}|Home_{c,t}, x_{c,t}] = 0$, where *Home* is their instrument defined as distance to place of birth of the candidate. In both first and second stage regression the author controls for county and time fixed effects and for deviation from normal vote shares in the election. Standard error are clustered by county to account for autocorrelation in the place of birth and residence of party candidates' across time.

Kinne (2013)

Q1: Extent of belonging to the same network on bilateral agreement probability.

Q2: Ideally it would be necessary to randomize the existence of previous agreements: the randomization should take place at the level of existing bilateral agreement, and existing third party agreements. In fact, the effect should be detected for both direct and indirect network connection.

Q3: To identify the effect the author constructs network variables that detect the extent of preference for bilateral agreement by evaluating: (1) the existence of previous agreement; (2) giving higher weights to countries that sign multiple agreements; (3) weight more partner that have similar agreements.

Q4: Stochastic actor-oriented model of network evolution, evaluate a dichotomous variable that assumes value 1 only a new connection is established at time τ. The variable is then used to generate indicator variables as in Q3. These represent explanatory variables within the estimation of a probit model on the signing of future agreements in general an by 4 other different categories: Commodities, Military, Science and Fisheries.

Jha (2013)

Q1: Long-term persistence of institutional effect. In this context, the effects of trade on tolerance levels.

Q2: The ideal treatment would be to randomly assign incentives to inter-ethnic trade and evaluate its effect on inter-ethnic relationships.

Q3: The explanatory variable is constructed based on whether medieval towns were trading ports with substantive evidence of oversea trade from historical evidence. To add for medieval era natural harbors, water bodies likely to be natural harbors within 10km of the modern Indian coast were included in the analysis. Each unit was than matched with its British era Indian district. The specification of the model than includes this explanatory variable in a statistical model regressing on the number of riots between 1850-1950. The specification controls for initial geographic conditions, medieval era measures of trade and human capital, colonial era fixed effects for native state, province, and timing of annexation.

Q4: The exclusion restriction is guaranteed by the assumption that: (1) The selection of location for medieval trade is uncorrelated with subsequent religious interaction. (2) Muslim traders have not chosen to trade at geographically similar ports for unobservable reasons. Standard error are clustered at Native state and annexation level. The statistical assumes uses both Negative binomial and Normal error distributions (OLS).

Flores Macias and Kreps (2013)

Q1: Presidential partisanship and the decision to finance war through specific taxation

Q2: Randomly assign president to party

Q3: Logistic regression: dependent variable is adoption of War tax (dichotomous) and explanatory variable of interest is whether the president belonged to party with pro-tax inclination. Controls: cost of conflict, inflation, debt/GDP, growth rate, divided government, electoral cycles, retaliation to attack, severity of war in given year

Q4: Logistic regression to estimate the effect of leaders partisan differences on war finance. Robust standard error to correct for heteroschedasticity.

Tomz and Weeks (2013)

Q1: Effect of recipient political regime on support for military action against that regime. The intent is to assess the effect of public opinion on the absence of wars among democracies.

Q2: Randomly assign political democratic regimes to country, and determine support for military action.

Q3: Randomly and independently assign four potential sources of peace: the political regime, alliance status, economic ties, and military power of the adversary and checked against support or opposition to a **preventive** military strike against the countrys nuclear facilities.

Q4: T-test of difference in averages of supporter for military action to democracy and non-democracies; military ally and non-ally; military strong country and not strong; and trade ally and non-ally.

Exercise 3

Part A - Description of the variables The histograms are provided below.

Figure 1: Distribution of X

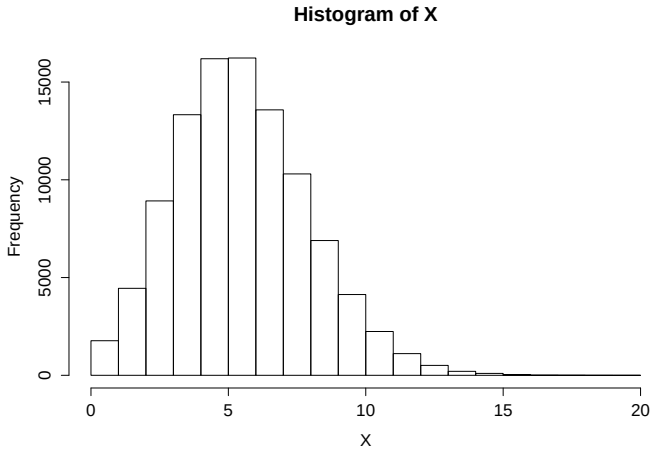
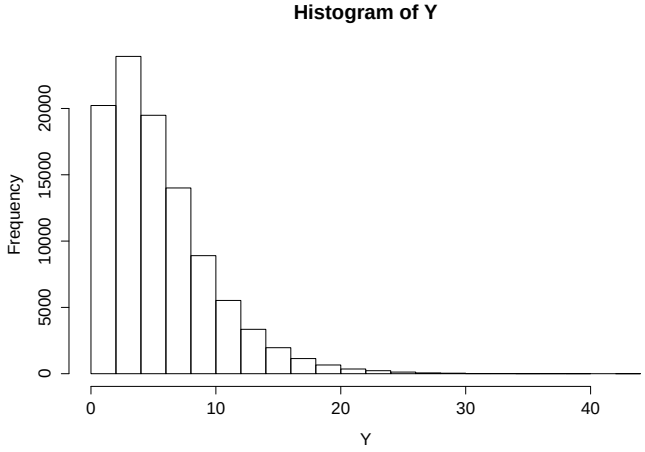
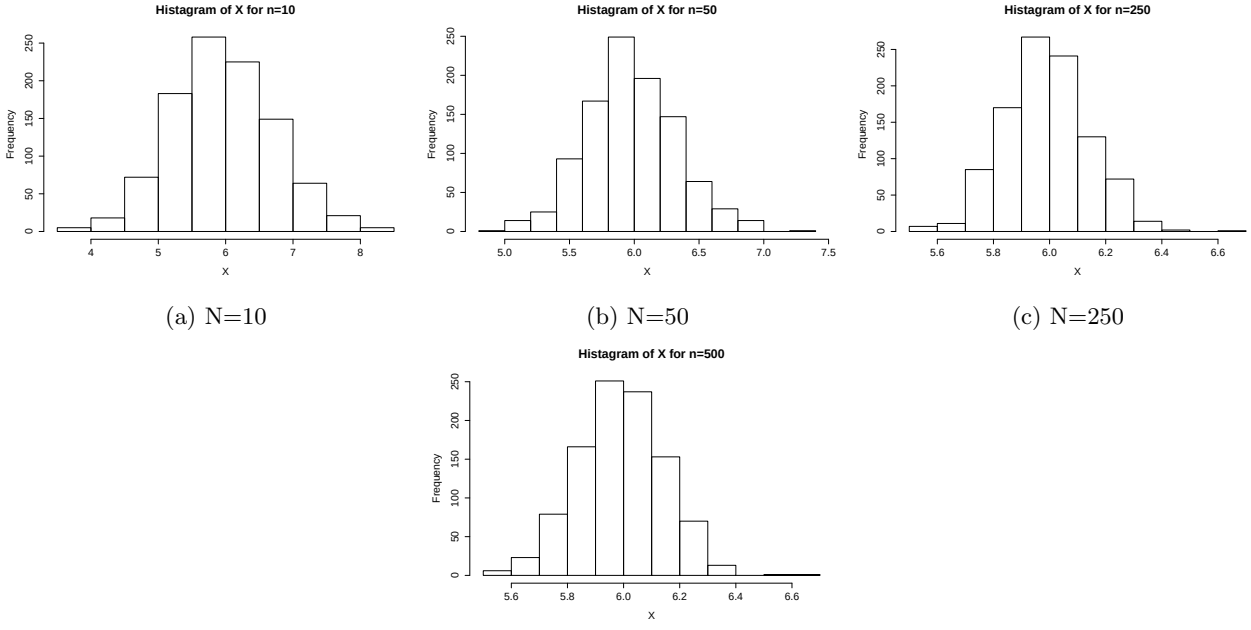


Figure 2: Distribution of Y



Notice that, with different extents, both distribution are skewed to the right. The histagrams show that the variance of X is larger than the one of Y due to the higher probability of extreme values in the support of X.



(a) N=10

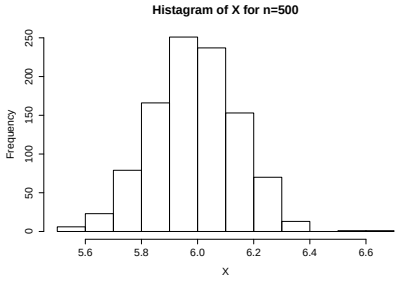
(b) N=50

(c) N=250

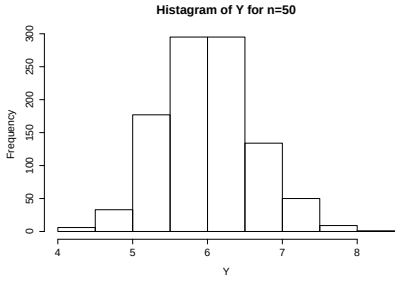
(d) N=500

Figure 3: Proof of the CLT for average of X

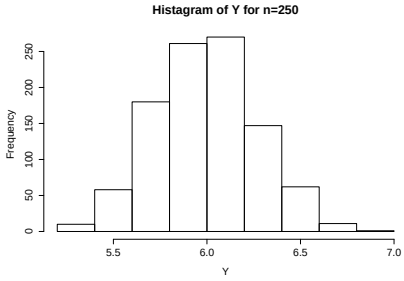
Part B - Simulations of convergence of the mean This section provides histograms for the 1000 simulations of the distribution of the sample mean for the two variables. Notice that indeed the distribution converges to a Normal distribution regardless to the initial distribution of the variables.



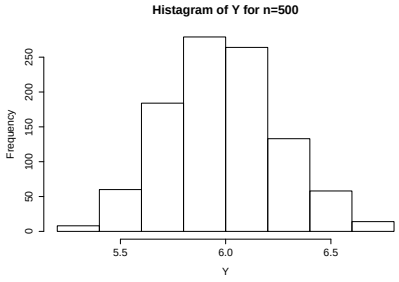
(a) N=10



(b) N=50



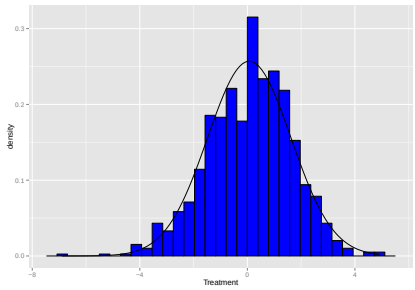
(c) N=250



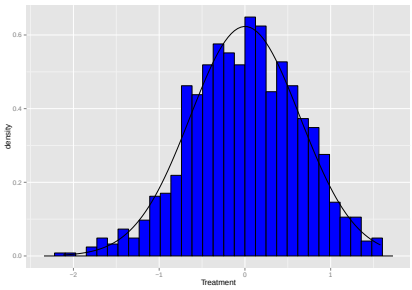
(d) N=500

Figure 4: Proof of the CLT for average of Y

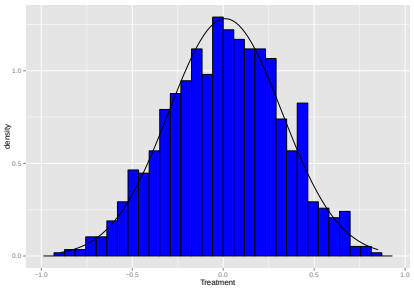
Part C - Simulations of randomized experiment This section reports the results from the simulation of randomized experiments ATE. The depicted variables are the ATE as estimated by the difference in mean outcome for the designed treatment and control groups. Notice that convergence to normality is pretty fast, as already with N=20 the distribution of the ATE looks normal. However, as expected, increases in the sample size reduce the spread of the distribution and increase the efficiency of the estimator.



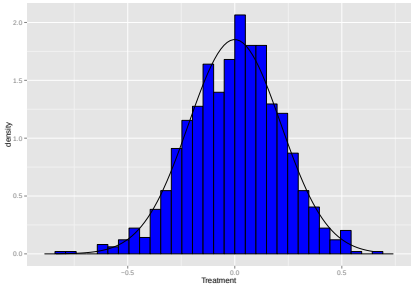
(a) N=20



(b) N=100



(c) N=500



(d) N=1000

Figure 5: Proof of the CLT for the ATE