

NAME: \_\_\_\_\_

**Experimental Research: Design, Analysis, and Interpretation**

W4368

Spring 2014

In-Class Midterm Exam

This exam reviews concepts from FEDAI Chapters 1-6. At the conclusion of this exam, please turn in this sheet along with your blue exam booklet.

Section I: Briefly define and state the significance of the following terms or phrases. Use formal notation to make your definitions as clear as possible.

1. Randomization inference
2. Covariate adjustment
3. Clustered random assignment
4. The assumption of monotonicity in the context of two-sided noncompliance

Section II: The following table was presented in Chapter 5. The results refer to the New Haven voter mobilization experiment, in which a random subset of the subject pool was assigned to be canvassed, but only some of those assigned to be canvassed were actually canvassed. The outcome is voter turnout.

**TABLE 5.2**

Voter turnout by experimental group, New Haven voter mobilization experiment

	Treatment group	Control group
Turnout rate among those contacted by canvassers	54.43 (395)	
Turnout rate among those not contacted by canvassers	36.48 (1,050)	37.54 (5,645)
Overall turnout rate	41.38 (1,445)	37.54 (5,645)

*Note:* Entries are percent voting, with number of observations in parentheses. Sample restricted to households containing a single registered voter.

1. Define a “Complier.”
2. Estimate the proportion of Compliers in the subject pool.
3. Show (with algebra) that under the assumption of non-interference and excludability, the CACE is identified in this application.
4. Are non-interference and excludability plausible in this example?
5. Estimate (by hand) the CACE. Provide a substantive interpretation of your estimate.

Section III. A recent experiment tested the effects of sending registered voters in New York City a (nonpartisan) postcard that encouraged them to donate to the candidate of their choice, on the grounds that small donations keep elected officials focused on important policy issues. Randomization was conducted within each of 5 blocks; blocks were created based on voters' expected probability of donating to campaigns in the future. Below is a block-by-block summary of the results. The first table displays the distribution of donations in control and treatment, by block. The next table displays the results of regressions, by block, of donations on treatment. The third table displays the results of inverse-probability weighted regression and the accompanying p-values derived from randomization inference under the sharp null hypothesis of no effect. The last table presents the results of an unweighted regression that controls for blocks. Based on your reading of the tables, answer the following questions:

1. Why would an (unweighted) regression of donations on treatment, ignoring blocking, be a biased estimator of the average treatment effect?
2. Interpret the third table's estimate of the ATE and the accompanying p-value.
3. The block-by-block regressions seem to suggest that the postcard has a significantly negative effect on donations in blocks 1 and 4. The p-values based on robust standard errors are below 0.05. Yet the overall estimate of the ATE using inverse-probability weights is positive. What do you think accounts for this apparent discrepancy?
4. Bonus: The weighted regression presented in table 3 produces an estimated ATE that is different from an unweighted regression (presented in table 4) that controls for blocks by using a dummy variable for each block (except one, which is the intercept). What, specifically, about the two regressions causes them to produce different estimates?

TABLE 1

	block	control_mean	control_sd	control_n	treatment_mean	treatment_sd	treatment_n
1	1	0.584	38.8	62315	0.000	0.00	300
2	2	0.804	41.9	30065	16.667	288.68	300
3	3	1.121	48.3	14788	0.833	14.43	300
4	4	2.031	73.2	7727	0.250	4.33	300
5	5	0.784	14.7	4198	8.417	144.34	300

TABLE 2

	amount_donated				
	Block 1	Block 2	Block 3	Block 4	Block 5
policy_postcard	-0.584*** (0.155)	15.863 (16.641)	-0.287 (0.922)	-1.781** (0.869)	7.632 (8.324)
Constant	0.584*** (0.155)	0.804*** (0.242)	1.121*** (0.397)	2.031** (0.832)	0.784*** (0.226)
N	62,615	30,365	15,088	8,027	4,498
R2	0.00000	0.001	0.00000	0.00002	0.002

\*p < .1; \*\*p < .05; \*\*\*p < .01

Robust Standard Errors in Parentheses

TABLE 3

Estimate from inverse probability weighted regression: 3.82112  
Two-sided p-value from randomization inference: 0.0317

TABLE 4

Linear regression

Number of obs = 120593  
F( 5,120587) = 1.22  
Prob > F = 0.2980  
R-squared = 0.0002  
Root MSE = 45.939

amount_donated	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
policy_postcard	4.165892	3.759707	1.11	0.268	-3.203073	11.53486
_lblock_2	.3585068	.3208548	1.12	0.264	-.2703635	.987377
_lblock_3	.4707502	.4237828	1.11	0.267	-.3598571	1.301358
_lblock_4	1.247713	.8301279	1.50	0.133	-.3793243	2.87475
_lblock_5	.4541807	.5669856	0.80	0.423	-.6571018	1.565463
_cons	.5612111	.155826	3.60	0.000	.2557946	.8666276