Experimental Research: Design, Analysis, and Interpretation
W4368
Spring 2013
Take-Home Midterm Exam

SOLUTION KEY (Draft)*

TA: Albert Fang

April 22, 2013

Contents

*I still need to double check the solution to Problem 2F (output is wacky). For Question 2, the solutions use a recalculated "persons" variable to define blocks, since the original variable "persons" is meaningless after student-specific datasets are generated.

Problem 1.

The FEDAI book presents results from a simplified version of the Clingingsmith et al. dataset. The actual dataset has multiple observations in each household, and random assignment of visas occurred at the level of the household (i.e., everyone in the household was entered in the lottery together, and everyone won or lost as a cluster). The data for this problem may be found in q1hajj_sub.dta. The assignment (success), treatment (hajj2006) and outcome variables (views) are otherwise the same as the dataset used in the book. As noted in Chapter 6, the study encountered two-sided noncompliance. Thus, when analyzing these data, bear in mind the complications of clustered assignment and two-sided noncompliance.

Problem 1, Part A

Conduct a randomization check using the pre-treatment covariates in the dataset (age, female, literate). Interpret the results.

General Solution:

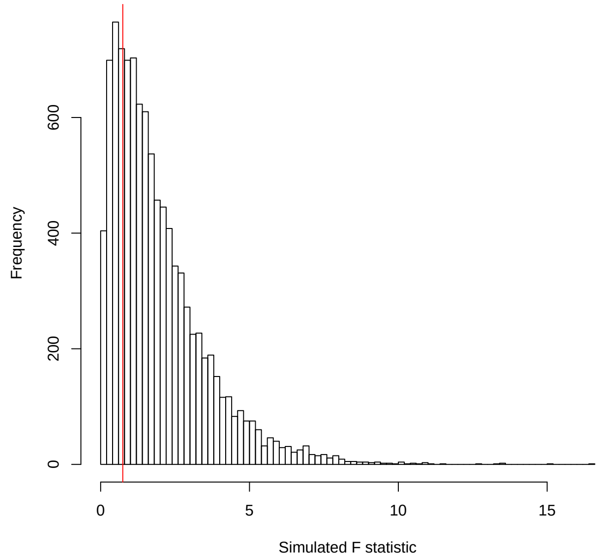
- Use randomization inference to test the null hypothesis that the covariates predict treatment assignment no better than would be expected by chance. The test statistic is the F statistic, which tells us the goodness of fit of a regression model (with all covariates) vs. a restricted model that contains only an intercept on the right hand side.
- First, regress treatment assignment (success) on three covariates (age, female, literate); grab F statistic.
- Simulate permutations of treatment assignment, accounting for clustered random assignment (clustvar=hhid). For each simulated vector of treatment assignments, regress the simulated treatment assignment vector on the three covariates, and save the F statistic from every iteration.
- Compute a p-value: the probability of obtaining an F statistic under the null hypothesis at least as large as the one obtained from the actual experiment. If the p-value is small (e.g., p=0.01), then this tells us the imbalance is greater than one would expect by chance (FEDAI, pp. 107-08).

Student-Specific Solutions:

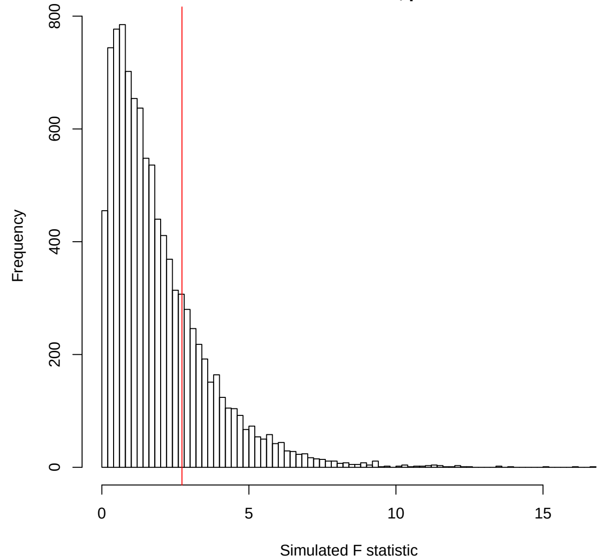
Student	f.stat	p.value
Anderson-Hill, Shayna	0.743	0.763
Ferrerosa-Young, Carolina	2.724	0.243
Foos, Florian	0.944	0.674
Khan, Sarah	2.170	0.345
Kirkland, Patricia Ann	0.859	0.708
Lazarev, Egor	0.526	0.832
Lozano, Andrea Patricia	1.280	0.528
Luby, Ryan Patrick	1.630	0.454
Marquez Pena, Javier	2.099	0.353
Moreno, Edgar Samuel	0.969	0.665
Pan, Yilin	1.482	0.497
Rink, Anselm Frieder	0.817	0.719
Sacramone-Lutz, Gabriella	1.389	0.507
Sharma, Kunaal	1.065	0.635
Snegovaya, Maria	1.625	0.467
Spry, Amber Denise	0.190	0.958
Tattersall, Laura	3.314	0.151
Warren, Shana	1.402	0.513
Zelizer, Adam Philip	5.493	0.040

Student-specific histograms of the simulated distribution of F statistics under the sharp null (# iterations: 10,000) are shown on the following pages. The observed F statistic from the experiment is shown in red.

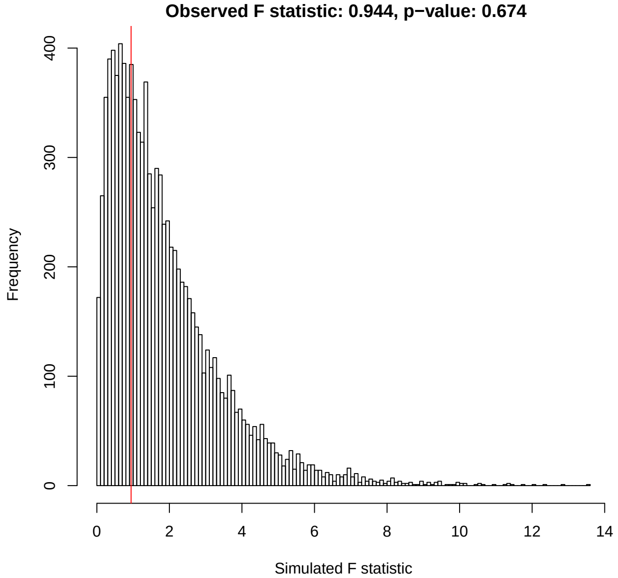
Student: Anderson−Hill, Shayna, UNI: saa2165
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.743, p−value: 0.763



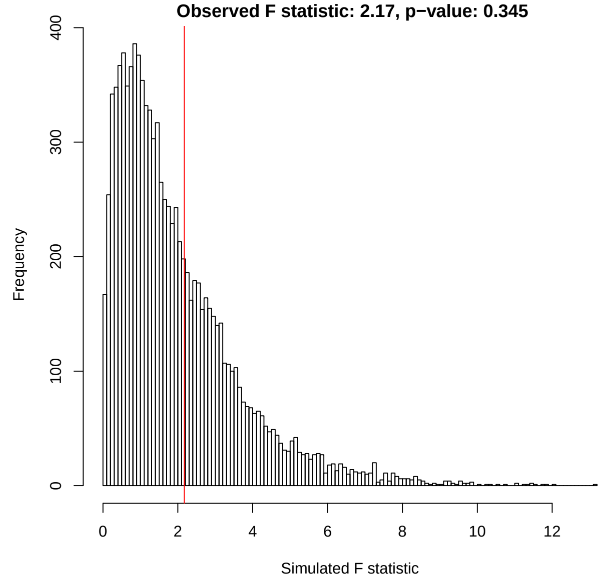
Student: Ferrerosa−Young, Carolina, UNI: cf2517
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.724, p−value: 0.244



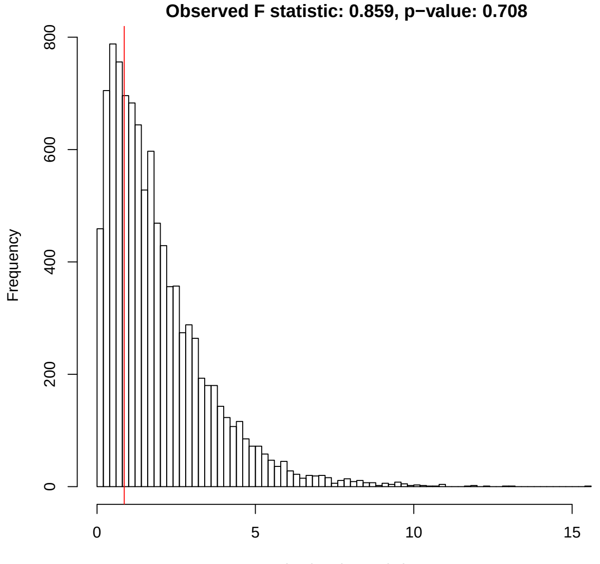
Student: Foos, Florian, UNI: ff2306
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.944, p−value: 0.674



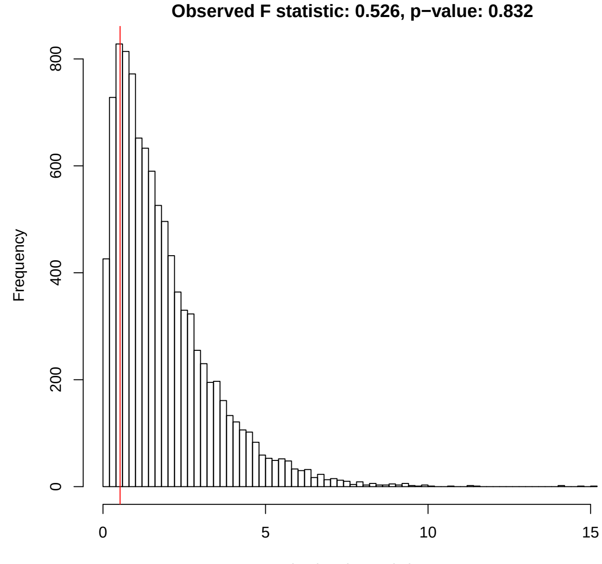
Student: Khan, Sarah, UNI: sk2947
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.17, p−value: 0.345

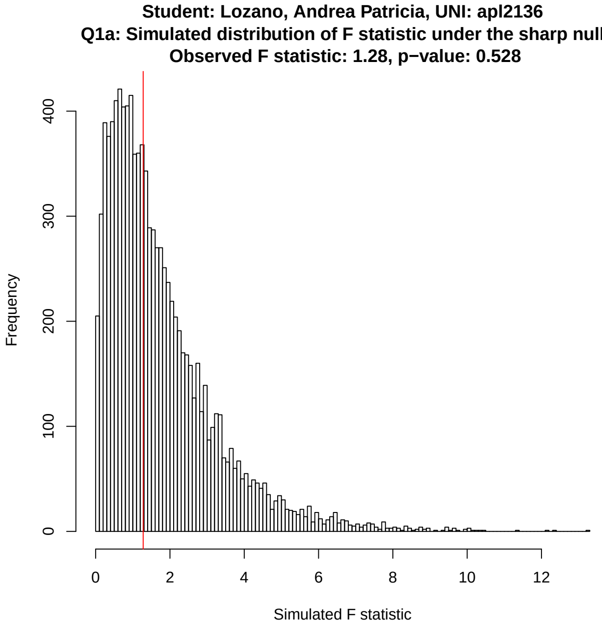


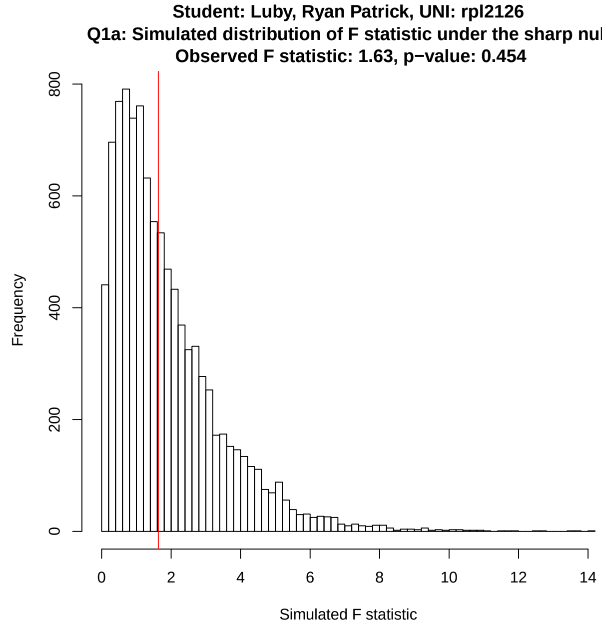
Student: Kirkland, Patricia Ann, UNI: pak2128
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.859, p−value: 0.708

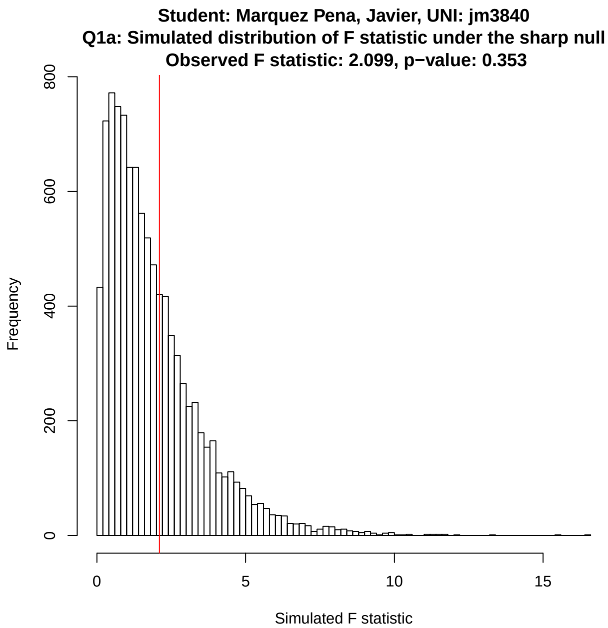


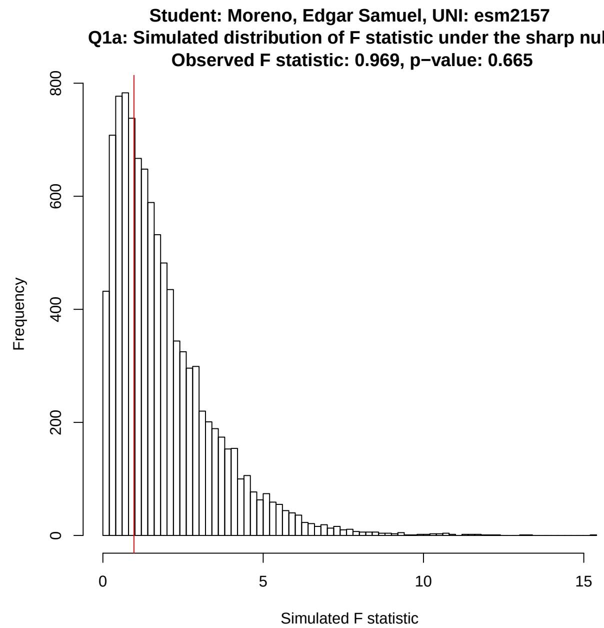
Student: Lazarev, Egor, UNI: el2666
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.526, p−value: 0.832

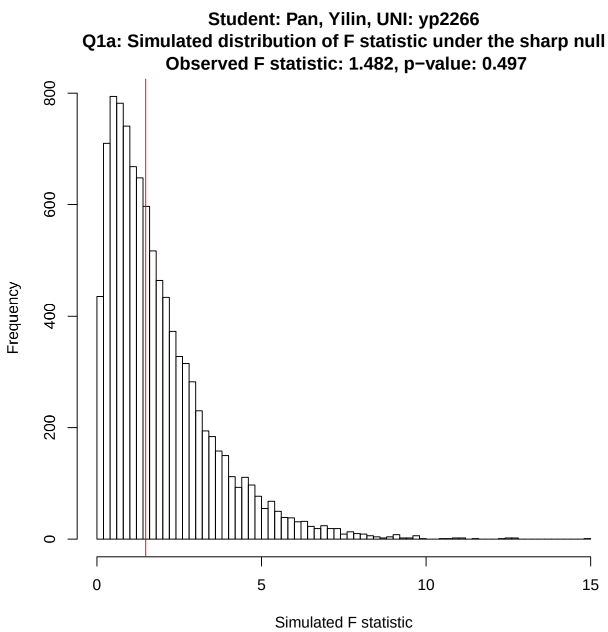


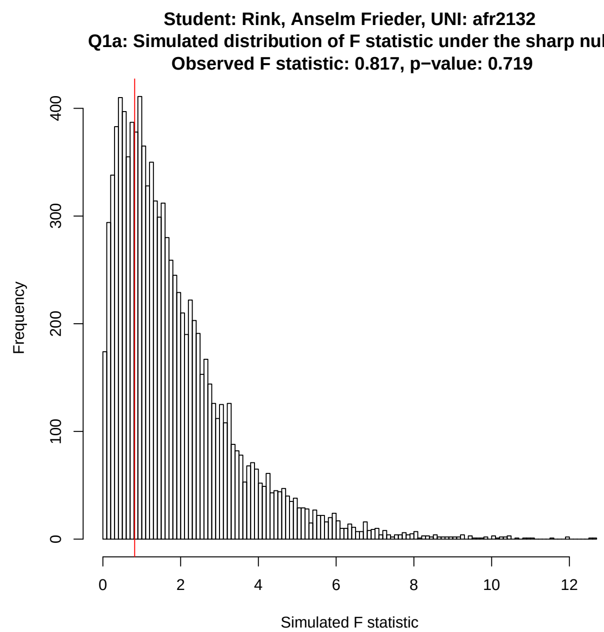

Student: Lozano, Andrea Patricia, UNI: apl2136
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.28, p−value: 0.528
Frequency
Simulated F statistic

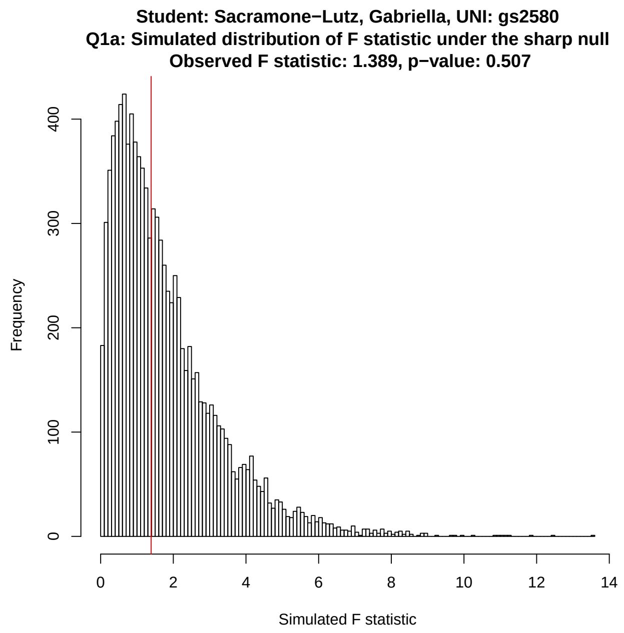

Student: Luby, Ryan Patrick, UNI: rpl2126
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.63, p−value: 0.454
Frequency
Simulated F statistic

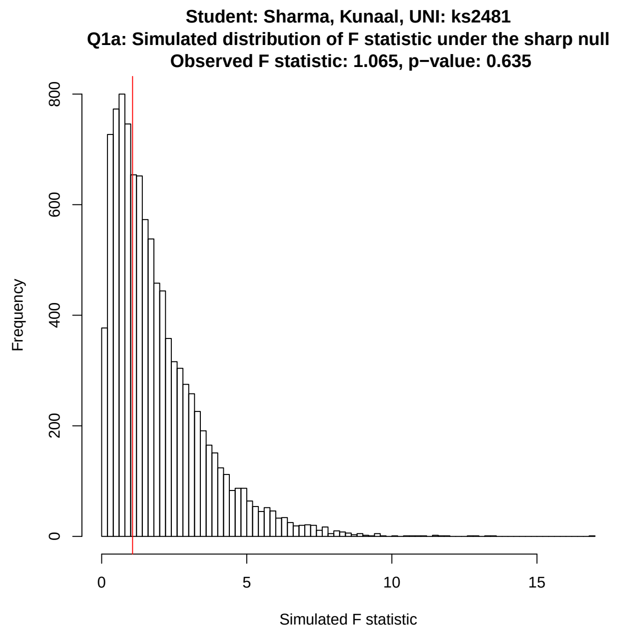

Student: Marquez Pena, Javier, UNI: jm3840
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.099, p−value: 0.353
Frequency
Simulated F statistic

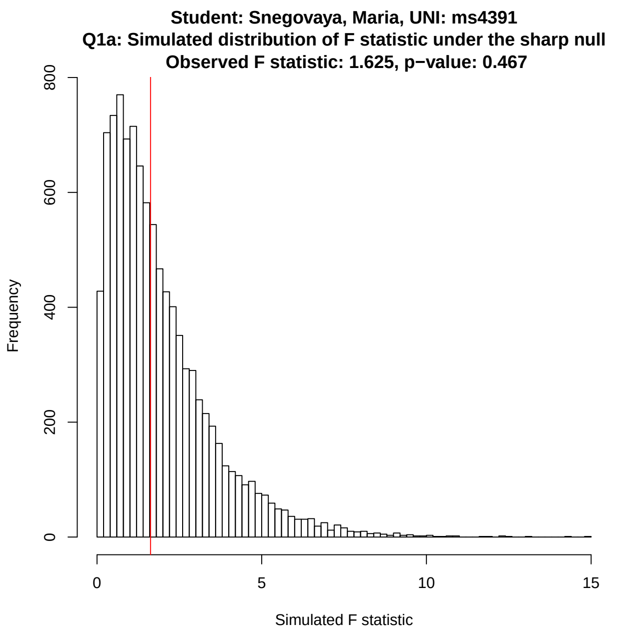

Student: Moreno, Edgar Samuel, UNI: esm2157
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.969, p−value: 0.665
Frequency
Simulated F statistic

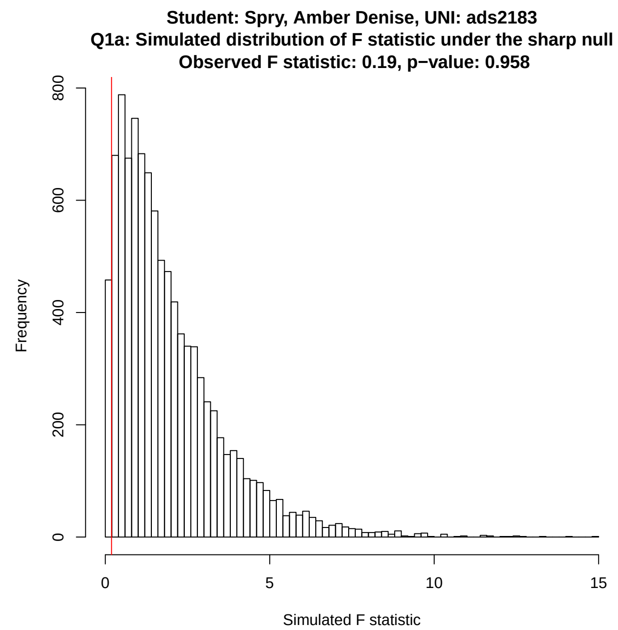

Student: Pan, Yilin, UNI: yp2266
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.482, p−value: 0.497
Frequency
Simulated F statistic

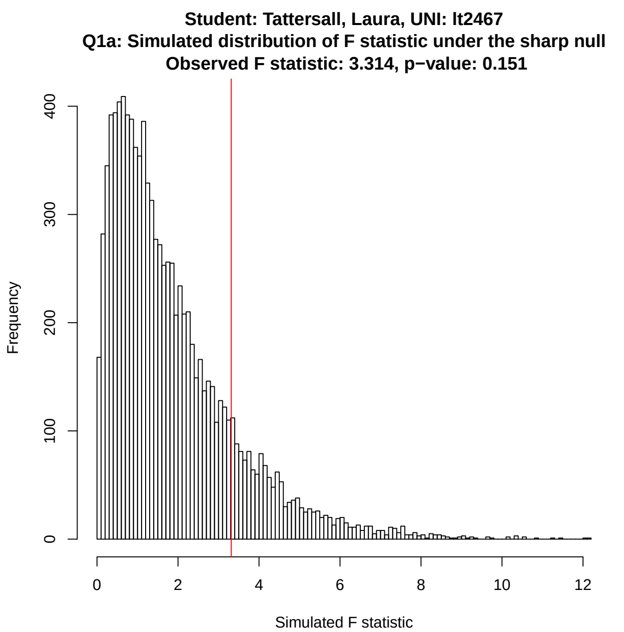

Student: Rink, Anselm Frieder, UNI: afr2132
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.817, p−value: 0.719
Frequency
Simulated F statistic

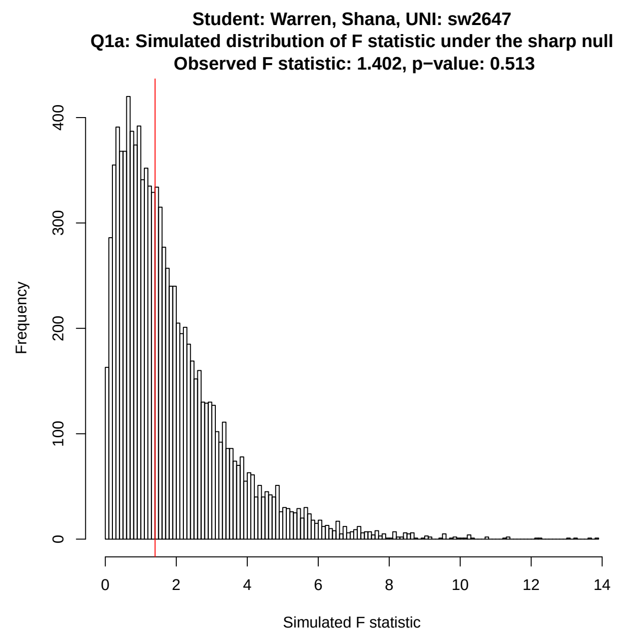

Student: Sacramone−Lutz, Gabriella, UNI: gs2580
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.389, p−value: 0.507
Frequency
Simulated F statistic

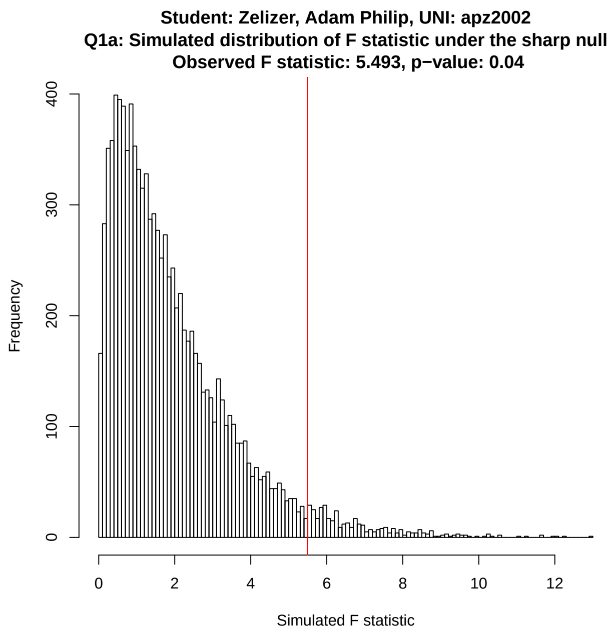

Student: Sharma, Kunaal, UNI: ks2481
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.065, p−value: 0.635
Frequency
Simulated F statistic


Student: Snegovaya, Maria, UNI: ms4391
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.625, p−value: 0.467
Frequency
Simulated F statistic


Student: Spry, Amber Denise, UNI: ads2183
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.19, p−value: 0.958
Frequency
Simulated F statistic


Student: Tattersall, Laura, UNI: lt2467
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 3.314, p−value: 0.151
Frequency
Simulated F statistic


Student: Warren, Shana, UNI: sw2647
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.402, p−value: 0.513
Frequency
Simulated F statistic


Student: Zelizer, Adam Philip, UNI: apz2002
Q1a: Simulated distribution of F statistic under the sharp null
Observed F statistic: 5.493, p−value: 0.04
Frequency
0
100
200
300
400
0
2
4
6
8
10
12
Simulated F statistic

Problem 1, Part B

Construct a table or graph to illustrate the ITT.

General Solution:

Estimate: $ITT = \mathbb{E}[Y_i(Z=1)] - \mathbb{E}[Y_i(Z=0)]$. The cell containing the ITT is noted in the table shell (similar to Table 6.2 in FEDAI) below.

	Treatment ($z=1$)		Control ($z=0$)			
	Mean	[N]	Mean	[N]	Difference	[N]
Treated ($d=1$)						
Untreated ($d=0$)						
Total					ITT	

Student-Specific Solutions:

Anderson-Hill, Shayna

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.292	[867]	3.152	[112]	-0.86	[979]
Untreated ($d=0$)	2	[5]	1.734	[597]	0.266	[602]
total	2.29	[872]	1.958	[709]	0.332	[1581]

Ferrerosa-Young, Carolina

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.373	[845]	2.72	[93]	-0.347	[938]
Untreated ($d=0$)	3	[3]	1.761	[640]	1.239	[643]
total	2.375	[848]	1.883	[733]	0.492	[1581]

Foos, Florian

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.35	[842]	2	[93]	0.35	[935]
Untreated ($d=0$)	1.8	[5]	1.827	[641]	-0.027	[646]
total	2.347	[847]	1.849	[734]	0.498	[1581]

Khan, Sarah

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.24	[853]	2.393	[89]	-0.153	[942]
Untreated ($d=0$)	3.25	[8]	1.777	[631]	1.473	[639]
total	2.25	[861]	1.853	[720]	0.397	[1581]

Kirkland, Patricia Ann

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.405	[843]	2.838	[99]	-0.433	[942]
Untreated ($d = 0$)	6.25	[4]	1.784	[635]	4.466	[639]
total	2.423	[847]	1.926	[734]	0.497	[1581]

Lazarev, Egor

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.198	[853]	2.368	[106]	-0.17	[959]
Untreated ($d = 0$)	2.556	[9]	1.705	[613]	0.851	[622]
total	2.202	[862]	1.803	[719]	0.399	[1581]

Lozano, Andrea Patricia

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.202	[803]	2.175	[114]	0.027	[917]
Untreated ($d = 0$)	3.8	[5]	1.774	[659]	2.026	[664]
total	2.212	[808]	1.833	[773]	0.379	[1581]

Luby, Ryan Patrick

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.253	[841]	2.914	[105]	-0.661	[946]
Untreated ($d = 0$)	3.25	[4]	1.721	[631]	1.529	[635]
total	2.258	[845]	1.891	[736]	0.367	[1581]

Marquez Pena, Javier

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.334	[854]	2.615	[96]	-0.281	[950]
Untreated ($d = 0$)	2	[7]	1.918	[624]	0.082	[631]
total	2.331	[861]	2.011	[720]	0.32	[1581]

Moreno, Edgar Samuel

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.514	[837]	2.728	[114]	-0.214	[951]
Untreated ($d = 0$)	3.286	[7]	1.724	[623]	1.562	[630]
total	2.52	[844]	1.879	[737]	0.641	[1581]

Pan, Yilin

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.261	[876]	2.644	[90]	-0.383	[966]
Untreated ($d=0$)	3.273	[11]	1.82	[604]	1.453	[615]
total	2.274	[887]	1.927	[694]	0.347	[1581]

Rink, Anselm Frieder

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.384	[812]	2.464	[97]	-0.08	[909]
Untreated ($d=0$)	1.6	[5]	1.775	[667]	-0.175	[672]
total	2.379	[817]	1.863	[764]	0.516	[1581]

Sacramone-Lutz, Gabriella

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.194	[836]	2.575	[120]	-0.381	[956]
Untreated ($d=0$)	4	[5]	1.615	[620]	2.385	[625]
total	2.205	[841]	1.77	[740]	0.435	[1581]

Sharma, Kunaal

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.556	[798]	2.876	[129]	-0.32	[927]
Untreated ($d=0$)	1.846	[13]	1.782	[641]	0.064	[654]
total	2.545	[811]	1.965	[770]	0.58	[1581]

Snegovaya, Maria

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.283	[817]	2.723	[112]	-0.44	[929]
Untreated ($d=0$)	2.25	[8]	1.806	[644]	0.444	[652]
total	2.282	[825]	1.942	[756]	0.34	[1581]

Spry, Amber Denise

	Treatment ($z=1$)		Control ($z=0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d=1$)	2.421	[822]	2.718	[124]	-0.297	[946]
Untreated ($d=0$)	1.5	[6]	1.836	[629]	-0.336	[635]
total	2.414	[828]	1.981	[753]	0.433	[1581]

Tattersall, Laura

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.448	[812]	2.603	[126]	-0.155	[938]
Untreated ($d = 0$)	3.375	[8]	1.587	[635]	1.788	[643]
total	2.457	[820]	1.756	[761]	0.701	[1581]

Warren, Shana

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.396	[847]	2.569	[109]	-0.173	[956]
Untreated ($d = 0$)	2.857	[7]	1.681	[618]	1.176	[625]
total	2.399	[854]	1.814	[727]	0.585	[1581]

Zelizer, Adam Philip

	Treatment ($z = 1$)		Control ($z = 0$)		Difference	Total
	$\overline{Y}$	N	$\overline{Y}$	N	$\overline{Y(1)} - \overline{Y(0)}$	N
Treated ($d = 1$)	2.425	[847]	2.153	[98]	0.272	[945]
Untreated ($d = 0$)	3	[7]	1.898	[629]	1.102	[636]
total	2.43	[854]	1.933	[727]	0.497	[1581]

Problem 1, Part C

Estimate the ITT with and without covariate adjustment. Provide a 95% confidence interval for the ITT. Test the sharp null hypothesis that the ITT is zero for all subjects.

<u>General Solution:</u>

Estimate the ITT using the difference-in-means estimator. Run a linear model without covariates (`Y ~ Z`), then with covariates (`Y ~ Z + X`). Inverse probability weights must be applied where weights equal the inverse of the probability of assignment to the observed assignment condition. Subject-specific probabilities of assignment to treatment can be calculated using the `ri` package using the `genprobexact` function – one must account for clustering (clustvar=hhid); then weights are computed as: `w <- Z/probs + (1-Z)/(1-probs)`.

To conduct a two-sided test of the sharp null that ITT = 0 for all subjects, generate all possible permutations of treatment assignment (or approximate using 10,000 simulated treatment assignment vectors) and estimate the ITT for each treatment permutation. Obtain the sampling distribution of the estimated ITT under the sharp null and calculate the probability of obtaining an ITT at least as large in absolute value as the one observed in the experiment.

We form a 95% confidence interval for the estimated ITT using the machinery of randomization inference (this is an approximate method for Rosenbaum's (2002) method of inverting the test statistic, per Gerber and Green 2012). We use the same procedure as randomization inference, but we set `ate=itt` when implementing the `genouts` function in the `ri` package. The interval has a 95% probability of bracketing the true ITT.

Because a clustered random assignment procedure is used, when estimating confidence intervals, we apply a correction that expands the width of the 95% confidence intervals by $\sqrt{(k-1)/(k-2)}$ where k is the number of clusters. The adjusted 95% confidence interval reports a lower bound that subtracts $\frac{\sqrt{(k-1)/(k-2)}}{2}$ from the unadjusted lower bound and reports an upper bound that adds $\frac{\sqrt{(k-1)/(k-2)}}{2}$ to the unadjusted upper bound.

<u>Student-Specific Solutions:</u>

		No Covariate Adjustment						With Covariate Adjustment					
			RI p-values			Adj. 95% CI			RI p-values			Adj. 95% CI	
Student	k	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB
Anderson-Hill, Shayna	769	0.332	0.084	0.040	0.960	-0.546	1.207	0.329	0.077	0.038	0.962	-0.538	1.197
Ferrerosa-Young, Carolina	769	0.492	0.011	0.005	0.995	-0.379	1.364	0.464	0.013	0.006	0.994	-0.404	1.325
Foos, Florian	779	0.498	0.003	0.001	0.999	-0.341	1.334	0.515	0.003	0.001	0.999	-0.332	1.352
Khan, Sarah	757	0.397	0.023	0.011	0.990	-0.456	1.238	0.437	0.012	0.005	0.995	-0.416	1.277
Kirkland, Patricia Ann	791	0.496	0.004	0.003	0.997	-0.342	1.338	0.498	0.004	0.002	0.998	-0.339	1.337
Lazarev, Egor	794	0.399	0.021	0.011	0.989	-0.439	1.238	0.394	0.022	0.011	0.989	-0.446	1.233
Lozano, Andrea Patricia	803	0.379	0.032	0.015	0.985	-0.475	1.219	0.402	0.019	0.008	0.992	-0.441	1.236
Luby, Ryan Patrick	797	0.367	0.044	0.022	0.978	-0.493	1.222	0.397	0.025	0.013	0.987	-0.448	1.245
Marquez Pena, Javier	770	0.320	0.083	0.041	0.959	-0.543	1.180	0.326	0.071	0.037	0.963	-0.532	1.180
Moreno, Edgar Samuel	775	0.641	0.000	0.000	1.000	-0.203	1.485	0.656	0.000	0.000	1.000	-0.189	1.498
Pan, Yilin	772	0.347	0.049	0.023	0.977	-0.502	1.189	0.363	0.036	0.017	0.983	-0.481	1.197
Rink, Anselm Frieder	773	0.517	0.005	0.003	0.997	-0.350	1.384	0.528	0.004	0.002	0.998	-0.332	1.393
Sacramone-Lutz, Gabriella	787	0.434	0.012	0.006	0.994	-0.395	1.273	0.418	0.015	0.008	0.992	-0.408	1.254
Sharma, Kunaal	771	0.580	0.002	0.001	1.000	-0.296	1.444	0.586	0.001	0.000	1.000	-0.276	1.444
Snegovaya, Maria	770	0.341	0.059	0.029	0.971	-0.512	1.194	0.319	0.069	0.035	0.965	-0.523	1.164
Spry, Amber Denise	777	0.433	0.020	0.009	0.991	-0.435	1.301	0.416	0.022	0.011	0.989	-0.442	1.277
Tattersall, Laura	786	0.702	0.000	0.000	1.000	-0.145	1.546	0.682	0.000	0.000	1.000	-0.159	1.521
Warren, Shana	781	0.585	0.001	0.000	1.000	-0.264	1.435	0.608	0.001	0.000	1.000	-0.236	1.452
Zelizer, Adam Philip	778	0.497	0.011	0.005	0.995	-0.386	1.363	0.454	0.018	0.008	0.992	-0.427	1.319

Problem 1, Part D

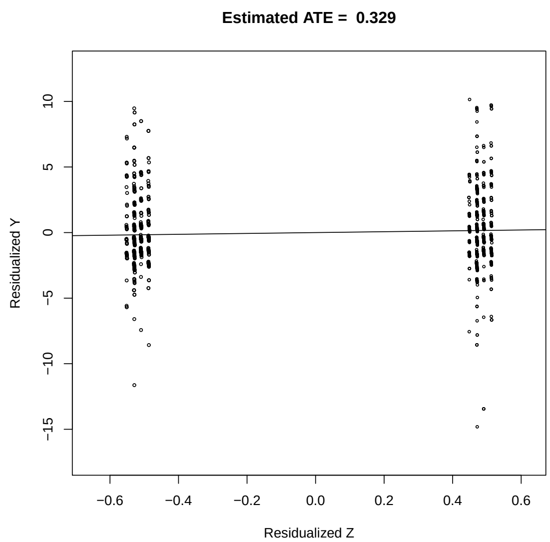
What are the pros and cons of estimating the ITT controlling for the pre-treatment covariates? Does controlling for covariates have a material effect on the results in part [b]? Present a res-res plot of the covariate-adjusted ITT (see Figure 4.3).

<u>General Solution:</u>
When we estimate the ITT controlling for pre-treatment covariates, we can improve the precision of our estimates if covariates are prognostic of potential outcomes. When the sample size is very small, covariate adjustment in a multiple regression framework generates biased point estimates of the treatment effect (i.e. Freedman's bias); the direction of this bias is unclear. Freedman's bias is not a concern in this application.

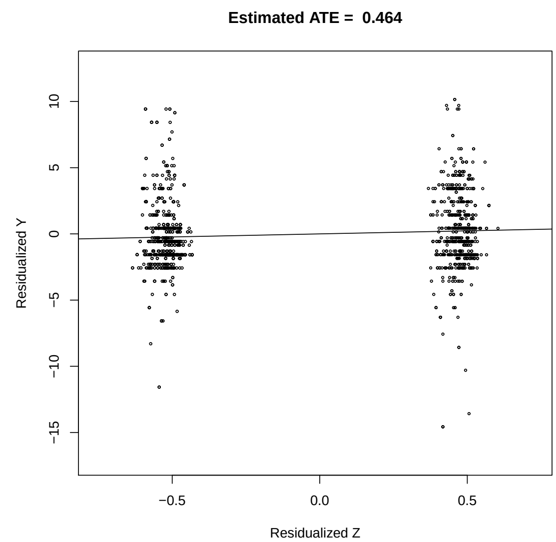
Controlling for covariates in the present application `[does/does not]` have a material effect on results from part [b]. (Note: Students should discuss, comparing specific results from parts [b] and [c]). In expectation over all possible random assignments, the estimator for the ITT that adjusts for covariates is unbiased. We might expect small differences in the estimates themselves to result from sampling variability (i.e., for a given randomization, the estimated ITT is drawn from the sampling distribution that is not at the mean). If we observe large differences, this may be the result of covariate imbalance by chance (assuming the randomization procedure and the ignorability assumption is valid).

<u>Student-Specific Solutions:</u>
Responses must discuss and compare ITT estimates without and with covariate adjustment.

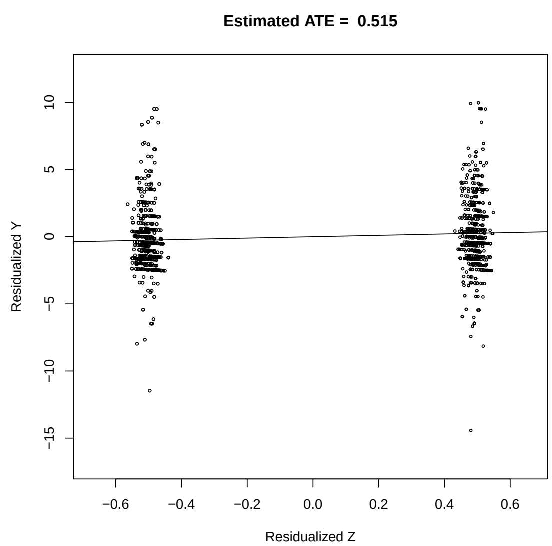
Res-res plots can be produced using the following function in the `ri` package: `resresplot(Y, Z, X=X, prob=probs, scale=1)`. The following pages contain res-res plots produced for each student. In the plot titles, "Estimated ATE" (the default output from the `resresplot` function) should read "Estimated ITT."



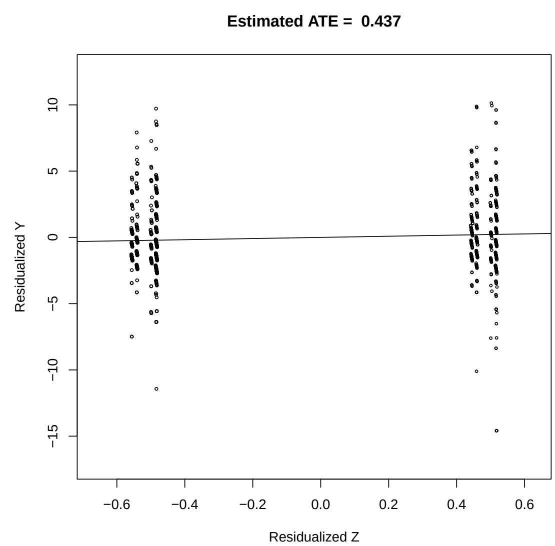
(a) Anderson-Hill, Shayna (UNI: saa2165)



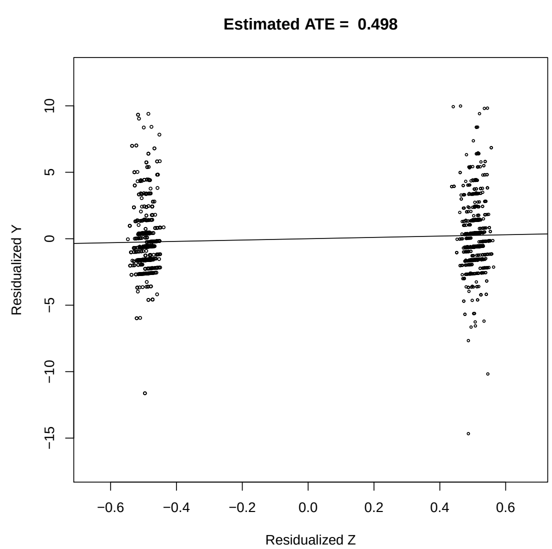
(b) Ferrerosa-Young, Carolina (UNI: cf2517)



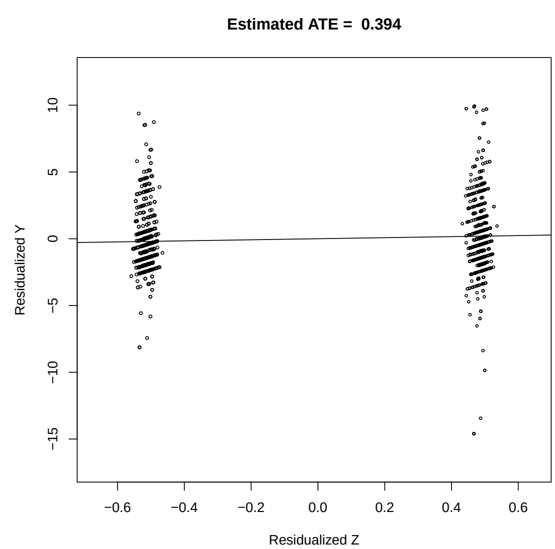
(c) Foos, Florian (UNI: ff2306)



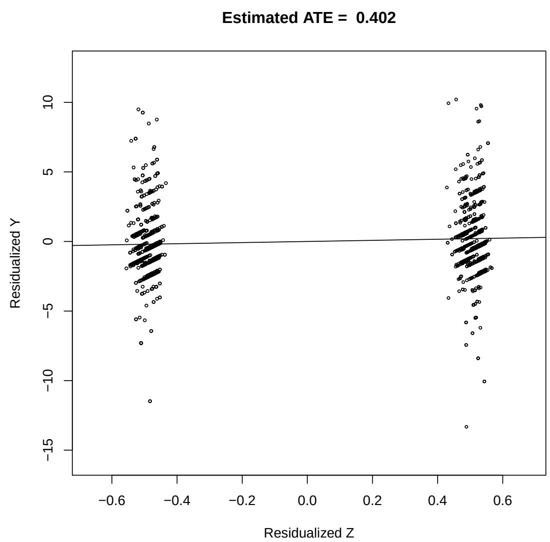
(d) Khan, Sarah (UNI: sk2947)



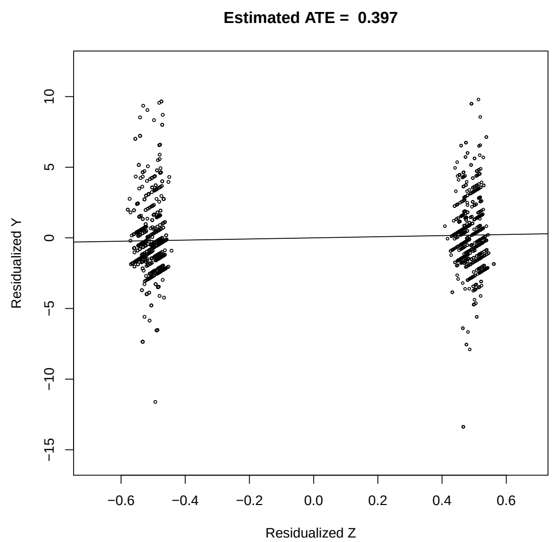
(e) Kirkland, Patricia Ann (UNI: pak2128)



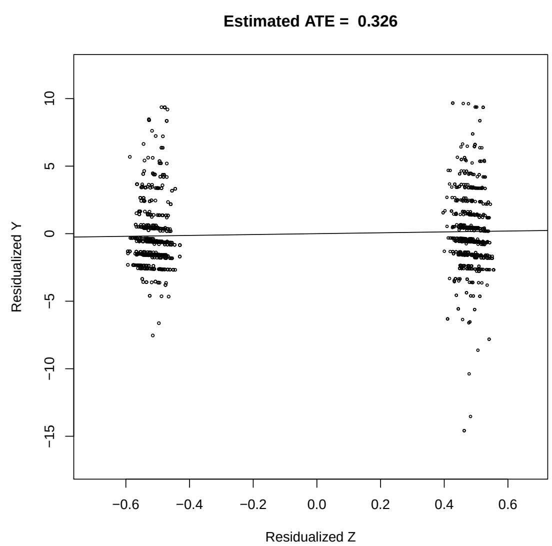
(f) Lazarev, Egor (UNI: el2666)



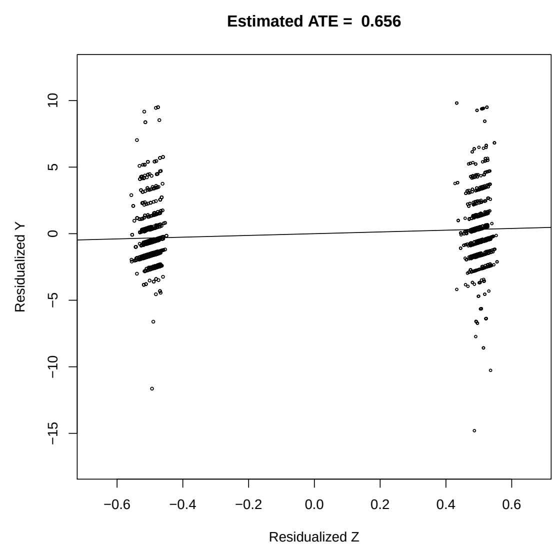
(g) Lozano, Andrea Patricia (UNI: apl2136)



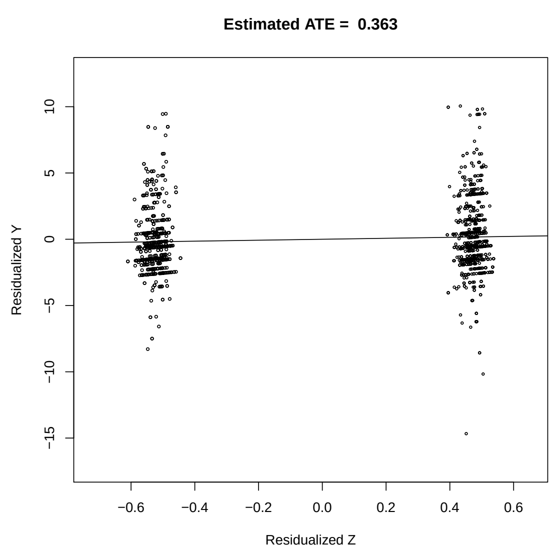
(h) Luby, Ryan Patrick (UNI: rpl2126)



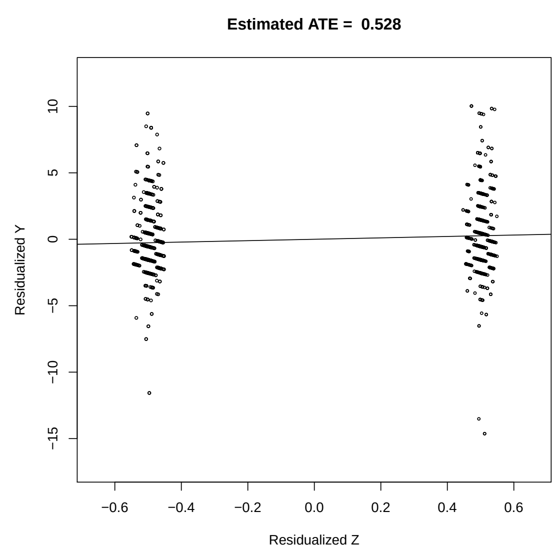
(i) Marquez Pena, Javier (UNI: jm3840)



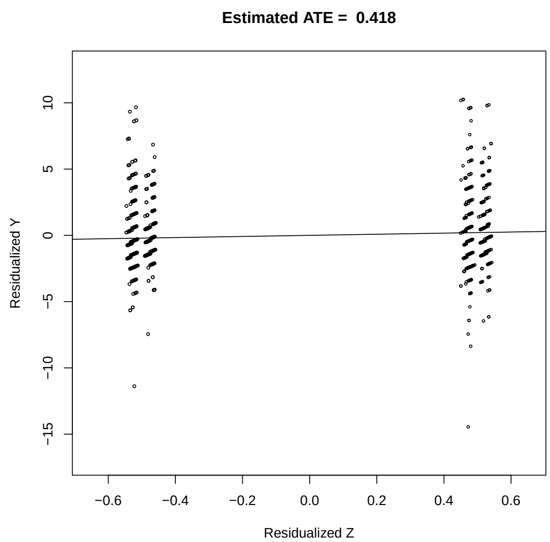
(j) Moreno, Edgar Samuel (UNI: esm2157)



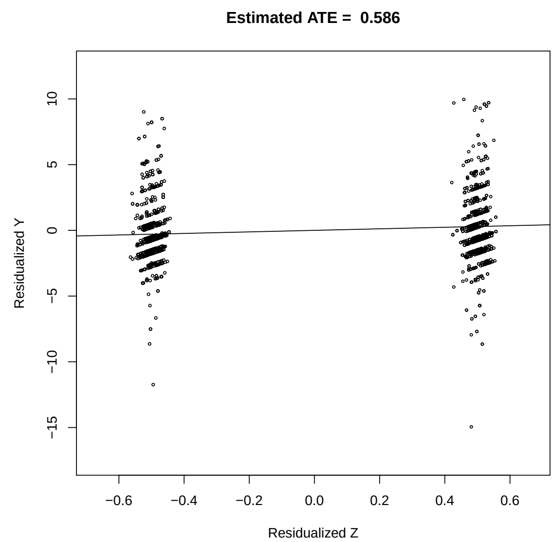
(k) Pan, Yilin (UNI: yp2266)



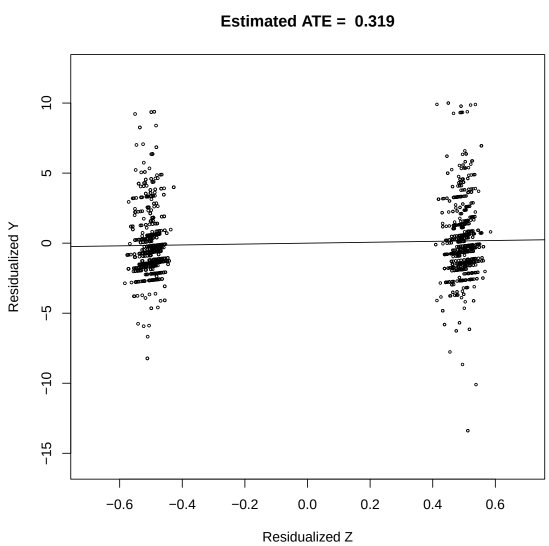
(l) Rink, Anselm Frieder (UNI: afr2132)



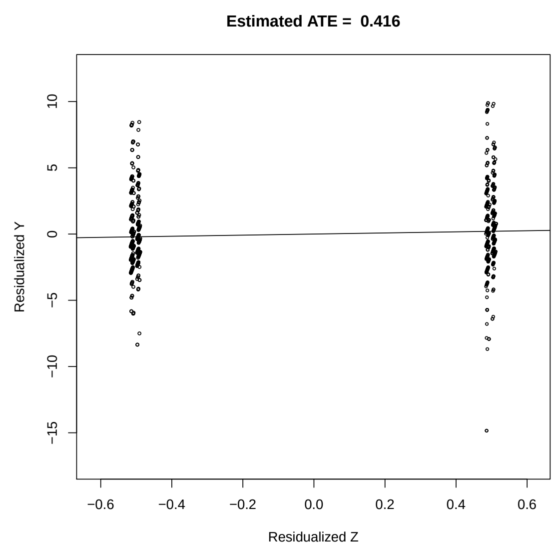
(m) Sacramone-Lutz, Gabriella (UNI: gs2580)



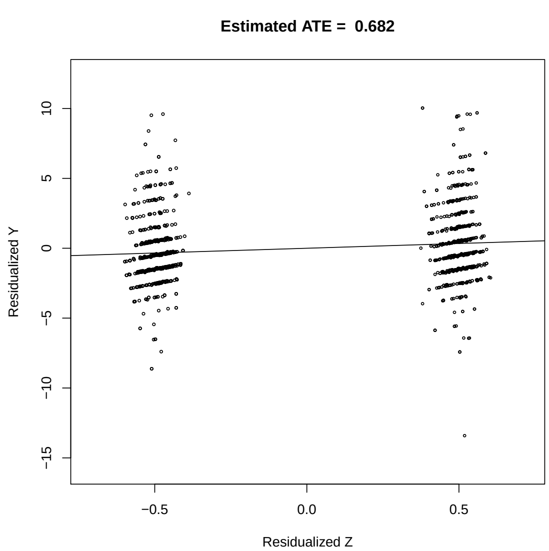
(n) Sharma, Kunaal (UNI: ks2481)



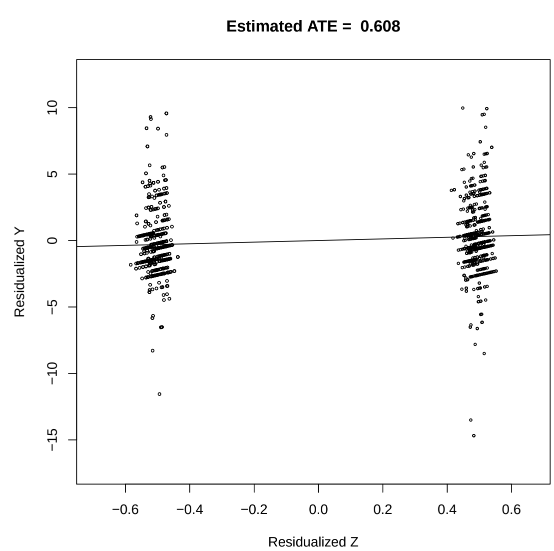
(o) Snegovaya, Maria (UNI: ms4391)



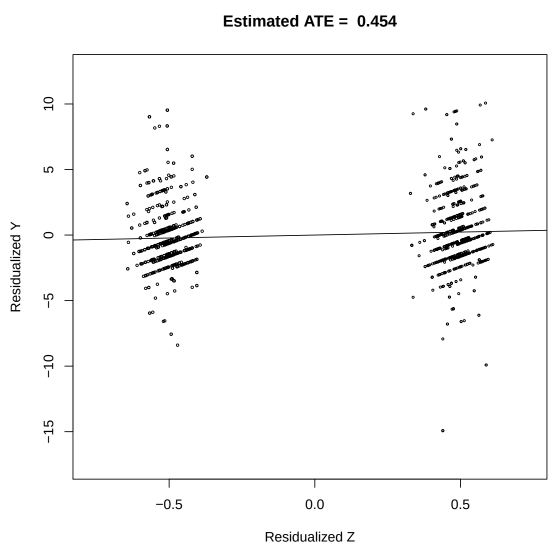
(p) Spry, Amber (UNI: ads2183)



(q) Tattersall, Laura (UNI: lt2467)



(r) Warren, Shana (UNI: sw2647)



(s) Zelizer, Adam Philip (UNI: apz2002)

Problem 1, Part E

Because cluster size varies, assess whether the results you found above change when you change the estimator of the ITT from difference-in-means to difference-in-totals.

General Solution:

The difference in totals estimator for the ITT is:

$$\widehat{ITT} = \frac{k_C + k_T}{N}\Big(\frac{\sum Y_i(1)|z_i = 1}{k_T} - \frac{\sum Y_i(0)|z_i = 0}{k_C}\Big)$$

We similarly adjust our estimates of the 95% confidence intervals to account for clustered randomization.

Student-Specific Solutions:

	No Covariate Adjustment						With Covariate Adjustment					
		RI p-values			Adj. 95% CI			RI p-values			Adj. 95% CI	
Student	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB
Anderson-Hill, Shayna	0.478	0.021	0.010	0.990	-0.430	1.369	0.327	0.077	0.039	0.961	-0.540	1.193
Ferrerosa-Young, Carolina	0.696	0.001	0.001	0.999	-0.218	1.610	0.460	0.014	0.006	0.994	-0.409	1.319
Foos, Florian	0.545	0.004	0.002	0.999	-0.330	1.416	0.513	0.002	0.001	0.999	-0.333	1.347
Khan, Sarah	0.529	0.006	0.003	0.998	-0.355	1.414	0.433	0.012	0.005	0.995	-0.421	1.269
Kirkland, Patricia Ann	0.428	0.033	0.016	0.984	-0.466	1.320	0.499	0.004	0.002	0.998	-0.338	1.334
Lazarev, Egor	0.509	0.005	0.002	0.998	-0.348	1.373	0.392	0.022	0.011	0.989	-0.448	1.229
Lozano, Andrea Patricia	0.322	0.100	0.050	0.951	-0.563	1.202	0.402	0.019	0.008	0.992	-0.441	1.234
Luby, Ryan Patrick	0.450	0.024	0.012	0.988	-0.441	1.337	0.394	0.025	0.013	0.987	-0.451	1.238
Marquez Pena, Javier	0.459	0.027	0.013	0.988	-0.456	1.363	0.323	0.072	0.037	0.963	-0.535	1.174
Moreno, Edgar Samuel	0.627	0.001	0.000	1.000	-0.255	1.492	0.655	0.000	0.000	1.000	-0.191	1.493
Pan, Yilin	0.566	0.003	0.001	0.999	-0.311	1.431	0.358	0.039	0.018	0.982	-0.485	1.190
Rink, Anselm Frieder	0.489	0.016	0.009	0.991	-0.396	1.385	0.528	0.004	0.002	0.998	-0.333	1.389
Sacramone-Lutz, Gabriella	0.451	0.015	0.007	0.993	-0.411	1.315	0.418	0.015	0.008	0.992	-0.408	1.250
Sharma, Kunaal	0.574	0.011	0.005	0.995	-0.359	1.495	0.585	0.001	0.000	1.000	-0.278	1.438
Snegovaya, Maria	0.360	0.075	0.038	0.963	-0.536	1.259	0.318	0.069	0.034	0.966	-0.524	1.159
Spry, Amber Denise	0.455	0.030	0.015	0.986	-0.456	1.359	0.416	0.022	0.011	0.989	-0.443	1.273
Tattersall, Laura	0.688	0.001	0.000	1.000	-0.198	1.571	0.677	0.000	0.000	1.000	-0.165	1.511
Warren, Shana	0.648	0.002	0.001	0.999	-0.246	1.538	0.606	0.001	0.000	1.000	-0.239	1.447
Zelizer, Adam Philip	0.556	0.011	0.006	0.995	-0.378	1.468	0.450	0.019	0.009	0.991	-0.431	1.313

Problem 1, Part F

Explain what the term "CACE" means in the context of this study.

<u>General Solution:</u>

The term "CACE" (complier average causal effect) refers to the average treatment effect of the hajj on views toward people from other countries among compliers – those who would go on the hajj if they win the lottery and would not go on the hajj if they did not win the lottery.

Problem 1, Part G

Explain what the monotonicity assumption means in the context of this study. Assuming monotonicity, estimate the share of Compliers, Never-Takers, and Always-Takers.

General Solution:

The monotonicity assumption states that $d_i(z=0) < d_i(z=1), \forall i \in N$ given treatment receipt $d = \{0,1\}$ and treatment assignment $z = \{0,1\}$ where 1=treatment, 0=control. In the context of this study, this assumption means that there are no individuals who are Defiers – that is, individuals who take up the treatment (i.e. go on the hajj) when assigned to control (i.e. losing the lottery) *and* who do not take up treatment (i.e. do not go on the hajj) when assigned to treatment (i.e. winning the lottery).

When we assume monotonicity, we can estimate the population shares of Compliers, Never-Takers, and Always-Takers. Recall that with a binary treatment variable, these three types are defined by the following relationships between $d_i(z)$ and z.

- Compliers: $d_i(0) = 0$ and $d_i(1) = 1$
- Always-Takers: $d_i(0) = 1$ and $d_i(1) = 1$
- Never-Takers: $d_i(0) = 0$ and $d_i(1) = 0$

Let the share of Compliers, Never-Takers, and Always-Takers be denoted:

$$\begin{aligned}
\alpha_C &= \Pr[d_i(0) = 0 \text{ and } d_i(1) = 1] = \frac{1}{N}\sum_{i=1}^{N} d_i(1)(1 - d_i(0)) = \Pr[\text{Complier}] \\
\alpha_A &= \Pr[d_i(0) = 1 \text{ and } d_i(1) = 1] = \frac{1}{N}\sum_{i=1}^{N} d_i(1)d_i(0) = \Pr[\text{Always-Taker}] \\
\alpha_N &= \Pr[d_i(0) = 0 \text{ and } d_i(1) = 0] = \frac{1}{N}\sum_{i=1}^{N} (1 - d_i(1))(1 - d_i(0)) = \Pr[\text{Never-Taker}]
\end{aligned}$$

and

$$1 = \alpha_C + \alpha_A + \alpha_N$$

Under two-sided noncompliance:

- Among those assigned to the treatment group ($z = 1$), subjects who do not take up treatment ($d = 0$) must be Never-Takers.
- Among those assigned to the control group ($z = 0$), subjects who do take up treatment ($d = 1$) must be Always-Takers.

Due to random assignment, in expectation the distribution of types is equal across treatment and control groups. So we can estimate:

$$\begin{aligned}
\hat{\alpha}_A &= \Pr[d = 1|z = 0] \\
\hat{\alpha}_N &= \Pr[d = 0|z = 1] \\
\hat{\alpha}_C &= 1 - \hat{\alpha}_A - \hat{\alpha}_N
\end{aligned}$$

Student-Specific Solutions:

Student	Pr(Complier)	Pr(Always-Taker)	Pr(Never-Taker)
Anderson-Hill, Shayna	0.836	0.158	0.006
Ferrerosa-Young, Carolina	0.870	0.127	0.004
Foos, Florian	0.867	0.127	0.006
Khan, Sarah	0.867	0.124	0.009
Kirkland, Patricia Ann	0.860	0.135	0.005
Lazarev, Egor	0.842	0.147	0.010
Lozano, Andrea Patricia	0.846	0.147	0.006
Luby, Ryan Patrick	0.853	0.143	0.005
Marquez Pena, Javier	0.859	0.133	0.008
Moreno, Edgar Samuel	0.837	0.155	0.008
Pan, Yilin	0.858	0.130	0.012
Rink, Anselm Frieder	0.867	0.127	0.006
Sacramone-Lutz, Gabriella	0.832	0.162	0.006
Sharma, Kunaal	0.816	0.168	0.016
Snegovaya, Maria	0.842	0.148	0.010
Spry, Amber Denise	0.828	0.165	0.007
Tattersall, Laura	0.825	0.166	0.010
Warren, Shana	0.842	0.150	0.008
Zelizer, Adam Philip	0.857	0.135	0.008

Problem 1, Part H

Estimate the CACE with and without covariate adjustment, and interpret the results.

<u>Student-Specific Solutions:</u>
CACE estimated using the IV estimator (`estlate` in the `ri` package, or the `ivreg` function in the `AER` package). P-values from randomization inference of the sharp null that the ITT = 0 for all subjects. Confidence intervals estimated using output from IV regression (assumes constant treatment effects and a normal sampling distribution) and are adjusted (for clustered randomization).

	No Covariate Adjustment						With Covariate Adjustment					
		RI p-values			Adj. 95% CI			RI p-values			Adj. 95% CI	
Student	$\widehat{CACE}$	Two-tailed	Greater	Lesser	LB	UB	$\widehat{CACE}$	Two-tailed	Greater	Lesser	LB	UB
Anderson-Hill, Shayna	0.398	0.084	0.040	0.960	-0.412	1.208	0.393	0.077	0.038	0.962	-0.411	1.198
Ferrerosa-Young, Carolina	0.566	0.011	0.005	0.995	-0.235	1.368	0.534	0.013	0.006	0.994	-0.265	1.334
Foos, Florian	0.575	0.003	0.001	0.999	-0.210	1.359	0.594	0.003	0.001	0.999	-0.189	1.376
Khan, Sarah	0.458	0.023	0.011	0.990	-0.326	1.242	0.503	0.012	0.005	0.995	-0.277	1.283
Kirkland, Patricia Ann	0.577	0.004	0.003	0.997	-0.210	1.364	0.579	0.004	0.002	0.998	-0.205	1.362
Lazarev, Egor	0.474	0.021	0.011	0.989	-0.320	1.268	0.468	0.022	0.011	0.989	-0.322	1.258
Lozano, Andrea Patricia	0.447	0.032	0.015	0.985	-0.346	1.241	0.474	0.019	0.008	0.992	-0.313	1.262
Luby, Ryan Patrick	0.430	0.044	0.022	0.978	-0.360	1.220	0.465	0.025	0.013	0.987	-0.318	1.248
Marquez Pena, Javier	0.373	0.083	0.041	0.959	-0.431	1.176	0.380	0.071	0.037	0.963	-0.420	1.180
Moreno, Edgar Samuel	0.766	0.000	0.000	1.000	-0.030	1.561	0.783	0.000	0.000	1.000	-0.008	1.574
Pan, Yilin	0.405	0.049	0.023	0.977	-0.389	1.199	0.423	0.036	0.017	0.983	-0.368	1.213
Rink, Anselm Frieder	0.596	0.005	0.003	0.997	-0.195	1.388	0.608	0.004	0.002	0.998	-0.179	1.396
Sacramone-Lutz, Gabriella	0.522	0.012	0.006	0.994	-0.268	1.312	0.503	0.015	0.008	0.992	-0.284	1.290
Sharma, Kunaal	0.710	0.002	0.001	1.000	-0.101	1.522	0.717	0.001	0.000	1.000	-0.088	1.522
Snegovaya, Maria	0.404	0.059	0.029	0.971	-0.397	1.206	0.379	0.069	0.035	0.965	-0.417	1.174
Spry, Amber Denise	0.523	0.020	0.009	0.991	-0.283	1.328	0.503	0.022	0.011	0.989	-0.297	1.303
Tattersall, Laura	0.851	0.000	0.000	1.000	0.057	1.645	0.826	0.000	0.000	1.000	0.038	1.615
Warren, Shana	0.695	0.001	0.000	1.000	-0.102	1.492	0.721	0.001	0.000	1.000	-0.072	1.514
Zelizer, Adam Philip	0.580	0.011	0.005	0.995	-0.224	1.384	0.530	0.018	0.008	0.992	-0.271	1.331

Problem 1, Part I

Another version of this dataset may be found in q1hajj.dta; this version includes observations for which outcomes are missing. Use this dataset to calculate extreme value bounds for the ITT.

<u>Student-Specific Solutions:</u>

Estimated extreme value bounds for the ITT, without and with covariate adjustment. The first four columns show the estimated extreme value bounds for the ITT estimated using the difference-in-means estimator. The next four columns show the estimated extreme value bounds for the ITT estimated using the difference-in-totals estimator.

	Difference-in-means estimator				Difference-in-totals estimator			
	Not Covariate Adjusted		Covariate Adjusted		Not Covariate Adjusted		Covariate Adjusted	
Student	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Anderson-Hill, Shayna	-0.013	0.815	0.025	0.861	0.180	1.012	0.025	0.853
Ferrerosa-Young, Carolina	0.133	0.831	0.180	0.874	0.302	1.008	0.179	0.867
Foos, Florian	0.179	0.948	0.178	0.950	0.155	0.924	0.177	0.947
Khan, Sarah	-0.055	0.958	-0.061	0.963	-0.020	0.993	-0.061	0.959
Kirkland, Patricia Ann	-0.225	0.571	-0.219	0.580	-0.304	0.489	-0.220	0.580
Lazarev, Egor	0.273	0.870	0.279	0.886	0.331	0.928	0.277	0.882
Lozano, Andrea Patricia	-0.045	0.641	-0.004	0.688	-0.101	0.584	-0.003	0.686
Luby, Ryan Patrick	-0.086	0.841	-0.096	0.846	-0.233	0.695	-0.095	0.845
Marquez Pena, Javier	0.079	0.649	0.096	0.661	0.136	0.705	0.096	0.658
Moreno, Edgar Samuel	-0.169	0.553	-0.179	0.543	-0.024	0.695	-0.178	0.541
Pan, Yilin	0.018	0.886	-0.032	0.830	-0.079	0.790	-0.034	0.830
Rink, Anselm Frieder	0.326	0.886	0.305	0.875	0.209	0.764	0.306	0.878
Sacramone-Lutz, Gabriella	0.062	0.709	0.044	0.695	0.170	0.819	0.044	0.692
Sharma, Kunaal	-0.059	0.781	-0.048	0.789	-0.059	0.781	-0.048	0.788
Snegovaya, Maria	0.120	0.841	0.094	0.818	0.211	0.930	0.094	0.811
Spry, Amber Denise	0.224	0.943	0.231	0.947	0.047	0.770	0.231	0.949
Tattersall, Laura	0.082	0.562	0.050	0.538	0.068	0.547	0.050	0.537
Warren, Shana	0.178	0.803	0.155	0.785	0.148	0.773	0.155	0.783
Zelizer, Adam Philip	0.213	0.785	0.203	0.778	0.234	0.806	0.203	0.776

Problem 1, Part J

Does missingness appear to be related to treatment assignment? Suppose the missingness rates were identical for the assigned treatment and control groups; under monotonicity, would the upper trimming bound and lower trimming bound be identical?

<u>General Solution:</u>
We can apply the logic of the randomization check (from Ch. 4) to conduct a formal statistical test of the null of random missingness (i.e. missingness is unrelated to treatment). We would regress observed missingness r_i on the experimental assignment and obtain an F-test statistic, and then we would compare this F statistic to the sampling distribution of F statistics generated using randomization inference.

[Student-specific solution: Insert results from statistical test here. Interpret.]

The main limitation of this method is that failure to reject the null of random missingness does not actually prove that missingness is unrelated to potential outcomes. (This holds even when we account for covariates prognostic of missingness.)

If missingness rates were identical for the assigned treatment and control groups; under monotonicity, the upper and lower trimming bounds would be identical because in expectation all of the reporters in both treatment and control are Always Reporters (i.e., there are no If-Treated Reporters in the treatment group). Thus per Eq. 7.21, $Q = 0$; the proportion of the Y_i values trimmed from the observed distribution in the treatment group is zero.

Problem 2.

In 2003, ACORN conducted an experiment in which 5,761 registered voters in Maricopa County were randomly assigned to be canvassed in advance of a municipal election. The data for this problem may be found in q2acorn.dta. Subjects resided in one-voter and two-voter households. Although the assignment was intended to take place at the individual level (see the variable treatment), ACORN decided that it was easier for canvassers to treat everyone in a household if any of its members were assigned to the treatment group. The effective assignment is the variable treat2; ACORN's procedure amounted to a blocked and clustered assignment, where blocks are defined by household size and clusters are all subjects living at the same address. The variable hhid gives the cluster identifier. The variable persons indicates the number of voters in each household. The outcome variable in this study is voter turnout, vote03. Previous votes, precinct, and age are included as covariates. ACORN's canvassers only reached some of the subjects they sought to canvass; the variable contact indicates that the treatment was actually administered. When analyzing these data, bear in mind the complications of blocked and clustered assignment as well as one-sided noncompliance.

Problem 2, Part A

Estimate the probability of treatment assignment for each block.

Anderson-Hill, Shayna

Block j (# Persons)	Pr(Assign to Treatment)
1	0.884
2	0.897
3	0.917
4	0.932
5	0.937
6	0.947
7	1

Ferrerosa-Young, Carolina

Block j (# Persons)	Pr(Assign to Treatment)
1	0.877
2	0.888
3	0.909
4	0.947
5	0.979
6	0.96
7	1
8	1

Foos, Florian

Block j (# Persons)	Pr(Assign to Treatment)
1	0.881
2	0.894
3	0.927
4	0.959
5	0.927
6	0.8
7	1
8	1

Khan, Sarah

Block j (# Persons)	Pr(Assign to Treatment)
1	0.884
2	0.893
3	0.928
4	0.946
5	0.97
6	1
7	1
9	1

Kirkland, Patricia Ann

Block j (# Persons)	Pr(Assign to Treatment)
1	0.88
2	0.898
3	0.907
4	0.956
5	0.979
6	0.933
7	1
8	1

Lazarev, Egor

Block j (# Persons)	Pr(Assign to Treatment)
1	0.884
2	0.894
3	0.911
4	0.957
5	0.944
6	0.933
7	1
8	1
9	1

Lozano, Andrea Patricia

Block j (# Persons)	Pr(Assign to Treatment)
1	0.877
2	0.882
3	0.893
4	0.957
5	1
6	1
7	1
10	1

Luby, Ryan Patrick

Block j (# Persons)	Pr(Assign to Treatment)
1	0.871
2	0.909
3	0.928
4	0.929
5	0.863
6	0.944
7	1
8	1

Marquez Pena, Javier

Block j (# Persons)	Pr(Assign to Treatment)
1	0.88
2	0.886
3	0.94
4	0.924
5	0.98
6	1
7	1

Moreno, Edgar Samuel

Block j (# Persons)	Pr(Assign to Treatment)
1	0.881
2	0.892
3	0.93
4	0.959
5	0.95
6	0.929
7	1

Pan, Yilin

Block j (# Persons)	Pr(Assign to Treatment)
1	0.866
2	0.919
3	0.917
4	0.95
5	0.887
6	1
7	1

Rink, Anselm Frieder

Block j (# Persons)	Pr(Assign to Treatment)
1	0.881
2	0.889
3	0.921
4	0.969
5	0.94
6	1
7	1
8	1

Sacramone-Lutz, Gabriella

Block j (# Persons)	Pr(Assign to Treatment)
1	0.864
2	0.907
3	0.916
4	0.964
5	0.949
6	0.941
7	0.8
9	1

Sharma, Kunaal

Block j (# Persons)	Pr(Assign to Treatment)
1	0.868
2	0.9
3	0.899
4	0.953
5	0.948
6	1
7	1

Snegovaya, Maria

Block j (# Persons)	Pr(Assign to Treatment)
1	0.878
2	0.896
3	0.927
4	0.932
5	0.933
6	0.933
7	1
8	1
9	1

Spry, Amber Denise

Block j (# Persons)	Pr(Assign to Treatment)
1	0.879
2	0.89
3	0.918
4	0.935
5	0.96
6	1
7	1

Tattersall, Laura

Block j (# Persons)	Pr(Assign to Treatment)
1	0.888
2	0.896
3	0.896
4	0.946
5	0.967
6	0.85
7	1
8	1
9	1
10	1

Warren, Shana

Block j (# Persons)	Pr(Assign to Treatment)
1	0.882
2	0.901
3	0.895
4	0.948
5	0.889
6	1
7	1

Zelizer, Adam Philip	
Block j (# Persons)	Pr(Assign to Treatment)
1	0.877
2	0.904
3	0.906
4	0.907
5	0.938
6	0.938
7	1
8	1

Problem 2, Part B

Construct a table showing the relationship between assigned treatment and voter turnout for each size household. Use this table to calculate the ITT for each block, and interpret the results.

Anderson-Hill, Shayna

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.182	1327	0.092	174	0.09	1501
2	0.172	1590	0.066	182	0.106	1772
3	0.163	1230	0.036	111	0.127	1341
4	0.248	656	0.083	48	0.165	704
5	0.217	295	0	20	0.217	315
6	0.25	108	0	6	0.25	114
7	0	14	NaN	0	NaN	14

Ferrerosa-Young, Carolina

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.168	1324	0.065	185	0.104	1509
2	0.171	1650	0.077	208	0.094	1858
3	0.16	1230	0.049	123	0.111	1353
4	0.214	576	0	32	0.214	608
5	0.204	235	0	5	0.204	240
6	0.215	144	0	6	0.215	150
7	0.286	35	NaN	0	NaN	35
8	0.625	8	NaN	0	NaN	8

Foos, Florian

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.169	1367	0.086	185	0.082	1552
2	0.177	1604	0.063	190	0.114	1794
3	0.164	1212	0.115	96	0.05	1308
4	0.189	656	0	28	0.189	684
5	0.137	255	0.25	20	-0.113	275
6	0.264	72	0	18	0.264	90
7	0.262	42	NaN	0	NaN	42
8	0.5	16	NaN	0	NaN	16

Khan, Sarah

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.157	1283	0.065	168	0.092	1451
2	0.172	1760	0.09	210	0.081	1970
3	0.19	1125	0.069	87	0.121	1212
4	0.188	628	0	36	0.188	664
5	0.215	325	0.1	10	0.115	335
6	0.283	60	NaN	0	NaN	60
7	0.119	42	NaN	0	NaN	42
9	0.333	27	NaN	0	NaN	27

Kirkland, Patricia Ann

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.165	1323	0.072	181	0.093	1504
2	0.176	1764	0.08	200	0.096	1964
3	0.169	1086	0.009	111	0.16	1197
4	0.218	696	0.094	32	0.125	728
5	0.217	230	0	5	0.217	235
6	0.202	84	0	6	0.202	90
7	0.4	35	NaN	0	NaN	35
8	0.375	8	NaN	0	NaN	8

Lazarev, Egor

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.169	1324	0.08	174	0.089	1498
2	0.184	1716	0.098	204	0.086	1920
3	0.163	1161	0.088	114	0.075	1275
4	0.163	620	0	28	0.163	648
5	0.212	255	0	15	0.212	270
6	0.298	84	0	6	0.298	90
7	0	35	NaN	0	NaN	35
8	1	16	NaN	0	NaN	16
9	1	9	NaN	0	NaN	9

Lozano, Andrea Patricia

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.169	1295	0.077	182	0.092	1477
2	0.185	1652	0.09	222	0.095	1874
3	0.19	1173	0.043	141	0.148	1314
4	0.161	628	0	28	0.161	656
5	0.269	305	NaN	0	NaN	305
6	0.367	90	NaN	0	NaN	90
7	0.314	35	NaN	0	NaN	35
10	0	10	NaN	0	NaN	10

Luby, Ryan Patrick

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.167	1279	0.069	189	0.099	1468
2	0.187	1790	0.033	180	0.153	1970
3	0.174	1200	0.129	93	0.045	1293
4	0.184	580	0.091	44	0.094	624
5	0.241	220	0	35	0.241	255
6	0.206	102	0	6	0.206	108
7	0.2	35	NaN	0	NaN	35
8	0	8	NaN	0	NaN	8

Marquez Pena, Javier

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.184	1343	0.049	184	0.135	1527
2	0.15	1670	0.103	214	0.047	1884
3	0.208	1221	0.154	78	0.054	1299
4	0.255	628	0	52	0.255	680
5	0.158	240	0	5	0.158	245
6	0.155	84	NaN	0	NaN	84
7	0.095	42	NaN	0	NaN	42

Moreno, Edgar Samuel

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.179	1299	0.103	175	0.076	1474
2	0.181	1700	0.097	206	0.084	1906
3	0.17	1188	0.067	90	0.103	1278
4	0.143	656	0	28	0.143	684
5	0.207	285	0	15	0.207	300
6	0.231	78	0	6	0.231	84
7	0.286	35	NaN	0	NaN	35

Pan, Yilin

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.172	1342	0.101	207	0.071	1549
2	0.178	1792	0.089	158	0.089	1950
3	0.186	1128	0.029	102	0.157	1230
4	0.17	612	0	32	0.17	644
5	0.174	235	0	30	0.174	265
6	0.186	102	NaN	0	NaN	102
7	0.095	21	NaN	0	NaN	21

Rink, Anselm Frieder

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.182	1350	0.071	182	0.111	1532
2	0.196	1664	0.053	208	0.143	1872
3	0.155	1152	0.212	99	-0.057	1251
4	0.169	632	0	20	0.169	652
5	0.203	315	0	20	0.203	335
6	0.222	90	NaN	0	NaN	90
7	0	21	NaN	0	NaN	21
8	0.5	8	NaN	0	NaN	8

Sacramone-Lutz, Gabriella

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.167	1276	0.08	200	0.087	1476
2	0.182	1652	0.059	170	0.123	1822
3	0.211	1236	0.123	114	0.088	1350
4	0.253	648	0	24	0.253	672
5	0.139	280	0	15	0.139	295
6	0.198	96	0	6	0.198	102
7	0.357	28	0	7	0.357	35
9	0	9	NaN	0	NaN	9

Sharma, Kunaal

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.182	1287	0.087	196	0.095	1483
2	0.182	1678	0.054	186	0.129	1864
3	0.148	1152	0.07	129	0.079	1281
4	0.192	652	0.25	32	-0.058	684
5	0.222	275	0	15	0.222	290
6	0.203	138	NaN	0	NaN	138
7	0.333	21	NaN	0	NaN	21

Snegovaya, Maria

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)]-\mathbb{E}[Y_i(0)]$	N_j
1	0.165	1343	0.081	186	0.084	1529
2	0.193	1694	0.117	196	0.076	1890
3	0.166	1173	0	93	0.166	1266
4	0.198	660	0.042	48	0.157	708
5	0.186	210	0	15	0.186	225
6	0.19	84	0	6	0.19	90
7	0.214	28	NaN	0	NaN	28
8	0	16	NaN	0	NaN	16
9	0	9	NaN	0	NaN	9

Spry, Amber Denise

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.177	1334	0.082	183	0.095	1517
2	0.17	1710	0.075	212	0.095	1922
3	0.193	1176	0.067	105	0.126	1281
4	0.181	636	0	44	0.181	680
5	0.362	240	0.5	10	-0.138	250
6	0.1	90	NaN	0	NaN	90
7	0	21	NaN	0	NaN	21

Tattersall, Laura

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.166	1320	0.072	166	0.094	1486
2	0.171	1730	0.095	200	0.076	1930
3	0.163	1137	0.045	132	0.117	1269
4	0.25	556	0	32	0.25	588
5	0.214	295	0	10	0.214	305
6	0.098	102	0	18	0.098	120
7	0.286	28	NaN	0	NaN	28
8	0	16	NaN	0	NaN	16
9	0	9	NaN	0	NaN	9
10	0	10	NaN	0	NaN	10

Warren, Shana

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.179	1330	0.101	178	0.078	1508
2	0.193	1656	0.082	182	0.111	1838
3	0.157	1200	0.043	141	0.114	1341
4	0.183	652	0.111	36	0.071	688
5	0.288	240	0.067	30	0.221	270
6	0.127	102	NaN	0	NaN	102
7	0	14	NaN	0	NaN	14

Zelizer, Adam Philip

Block j (# Persons)	$\mathbb{E}[Y_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[Y_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$	N_j
1	0.172	1344	0.069	189	0.103	1533
2	0.167	1672	0.067	178	0.1	1850
3	0.182	1218	0.087	126	0.095	1344
4	0.231	588	0	60	0.231	648
5	0.147	225	0	15	0.147	240
6	0.322	90	0	6	0.322	96
7	0.167	42	NaN	0	NaN	42
8	1	8	NaN	0	NaN	8

Problem 2, Part C

Pooling both blocks, estimate the overall ITT, and estimate a 95% confidence interval. Test the sharp null hypothesis of no ITT effect. Interpret the results.

		No Covariate Adjustment						With Covariate Adjustment					
			RI p-values			Adj. 95% CI			RI p-values			Adj. 95% CI	
Student	k	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB
Anderson-Hill, Shayna	3092.00	0.12	0.00	0.00	1.00	-0.43	0.67	0.11	0.00	0.00	1.00	-0.44	0.67
Ferrerosa-Young, Carolina	3114.00	0.12	0.00	0.00	1.00	-0.44	0.68	0.11	0.00	0.00	1.00	-0.45	0.67
Foos, Florian	3126.00	0.09	0.00	0.00	1.00	-0.47	0.64	0.08	0.00	0.00	1.00	-0.47	0.64
Khan, Sarah	3073.00	0.11	0.00	0.00	1.00	-0.45	0.66	0.10	0.00	0.00	1.00	-0.46	0.65
Kirkland, Patricia Ann	3129.00	0.12	0.00	0.00	1.00	-0.45	0.67	0.11	0.00	0.00	1.00	-0.45	0.67
Lazarev, Egor	3114.00	0.10	0.00	0.00	1.00	-0.45	0.65	0.09	0.00	0.00	1.00	-0.46	0.64
Lozano, Andrea Patricia	3016.00	0.12	0.00	0.00	1.00	-0.44	0.66	0.11	0.00	0.00	1.00	-0.44	0.66
Luby, Ryan Patrick	3109.00	0.11	0.00	0.00	1.00	-0.44	0.66	0.10	0.00	0.00	1.00	-0.45	0.65
Marquez Pena, Javier	3121.00	0.10	0.00	0.00	1.00	-0.46	0.66	0.10	0.00	0.00	1.00	-0.46	0.66
Moreno, Edgar Samuel	3098.00	0.10	0.00	0.00	1.00	-0.46	0.65	0.10	0.00	0.00	1.00	-0.46	0.65
Pan, Yilin	3148.00	0.11	0.00	0.00	1.00	-0.44	0.66	0.11	0.00	0.00	1.00	-0.44	0.66
Rink, Anselm Frieder	3115.00	0.10	0.00	0.00	1.00	-0.46	0.65	0.09	0.00	0.00	1.00	-0.47	0.65
Sacramone-Lutz, Gabriella	3086.00	0.12	0.00	0.00	1.00	-0.44	0.68	0.12	0.00	0.00	1.00	-0.44	0.67
Sharma, Kunaal	3071.00	0.09	0.00	0.00	1.00	-0.46	0.64	0.09	0.00	0.00	1.00	-0.47	0.64
Snegovaya, Maria	3133.00	0.11	0.00	0.00	1.00	-0.44	0.66	0.11	0.00	0.00	1.00	-0.45	0.66
Spry, Amber Denise	3125.00	0.10	0.00	0.00	1.00	-0.46	0.66	0.10	0.00	0.00	1.00	-0.46	0.65
Tattersall, Laura	3102.00	0.12	0.00	0.00	1.00	-0.44	0.67	0.11	0.00	0.00	1.00	-0.45	0.66
Warren, Shana	3100.00	0.10	0.00	0.00	1.00	-0.45	0.65	0.10	0.00	0.00	1.00	-0.45	0.65
Zelizer, Adam Philip	3132.00	0.12	0.00	0.00	1.00	-0.43	0.67	0.11	0.00	0.00	1.00	-0.44	0.66

Problem 2, Part D

Because cluster size varies, assess whether the results in part [c] change when you change the estimator of the ITT from difference-in-means to difference-in-totals.

		No Covariate Adjustment						With Covariate Adjustment					
			RI p-values			Adj. 95% CI			RI p-values			Adj. 95% CI	
Student	k	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB	$\widehat{ITT}$	Two-tailed	Greater	Lesser	LB	UB
Anderson-Hill, Shayna	3092.00	0.12	0.00	0.00	1.00	-0.43	0.67	0.11	0.00	0.00	1.00	-0.44	0.66
Ferrerosa-Young, Carolina	3114.00	0.12	0.00	0.00	1.00	-0.44	0.68	0.11	0.00	0.00	1.00	-0.45	0.66
Foos, Florian	3126.00	0.09	0.00	0.00	1.00	-0.47	0.64	0.08	0.00	0.00	1.00	-0.47	0.64
Khan, Sarah	3073.00	0.11	0.00	0.00	1.00	-0.45	0.66	0.09	0.00	0.00	1.00	-0.47	0.65
Kirkland, Patricia Ann	3129.00	0.12	0.00	0.00	1.00	-0.45	0.67	0.11	0.00	0.00	1.00	-0.45	0.67
Lazarev, Egor	3114.00	0.10	0.00	0.00	1.00	-0.45	0.65	0.09	0.00	0.00	1.00	-0.46	0.64
Lozano, Andrea Patricia	3016.00	0.12	0.00	0.00	1.00	-0.44	0.66	0.11	0.00	0.00	1.00	-0.44	0.66
Luby, Ryan Patrick	3109.00	0.11	0.00	0.00	1.00	-0.44	0.66	0.10	0.00	0.00	1.00	-0.45	0.65
Marquez Pena, Javier	3121.00	0.10	0.00	0.00	1.00	-0.46	0.66	0.10	0.00	0.00	1.00	-0.46	0.65
Moreno, Edgar Samuel	3098.00	0.10	0.00	0.00	1.00	-0.46	0.65	0.10	0.00	0.00	1.00	-0.46	0.65
Pan, Yilin	3148.00	0.11	0.00	0.00	1.00	-0.44	0.66	0.11	0.00	0.00	1.00	-0.44	0.66
Rink, Anselm Frieder	3115.00	0.10	0.00	0.00	1.00	-0.46	0.65	0.09	0.00	0.00	1.00	-0.47	0.65
Sacramone-Lutz, Gabriella	3086.00	0.12	0.00	0.00	1.00	-0.44	0.68	0.12	0.00	0.00	1.00	-0.44	0.67
Sharma, Kunaal	3071.00	0.09	0.00	0.00	1.00	-0.46	0.64	0.09	0.00	0.00	1.00	-0.47	0.64
Snegovaya, Maria	3133.00	0.11	0.00	0.00	1.00	-0.44	0.66	0.11	0.00	0.00	1.00	-0.45	0.66
Spry, Amber Denise	3125.00	0.10	0.00	0.00	1.00	-0.46	0.66	0.10	0.00	0.00	1.00	-0.46	0.65
Tattersall, Laura	3102.00	0.12	0.00	0.00	1.00	-0.44	0.67	0.11	0.00	0.00	1.00	-0.45	0.66
Warren, Shana	3100.00	0.10	0.00	0.00	1.00	-0.45	0.65	0.10	0.00	0.00	1.00	-0.45	0.65
Zelizer, Adam Philip	3132.00	0.12	0.00	0.00	1.00	-0.43	0.67	0.11	0.00	0.00	1.00	-0.44	0.66

Problem 2, Part E

Construct a table showing the relationship between assigned treatment and actual treatment for each size household. Use this table to calculate the ITT_D for each block, and interpret the results.

Anderson-Hill, Shayna

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.729	1327	0	174	0.729	1501
2	0.741	1590	0	182	0.741	1772
3	0.76	1230	0	111	0.76	1341
4	0.77	656	0	48	0.77	704
5	0.664	295	0	20	0.664	315
6	0.972	108	0	6	0.972	114
7	0.5	14	NaN	0	NaN	14

Ferrerosa-Young, Carolina

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.721	1324	0	185	0.721	1509
2	0.751	1650	0	208	0.751	1858
3	0.751	1230	0	123	0.751	1353
4	0.802	576	0	32	0.802	608
5	0.762	235	0	5	0.762	240
6	0.833	144	0	6	0.833	150
7	0.6	35	NaN	0	NaN	35
8	0.625	8	NaN	0	NaN	8

Foos, Florian

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.733	1367	0	185	0.733	1552
2	0.742	1604	0	190	0.742	1794
3	0.775	1212	0	96	0.775	1308
4	0.78	656	0	28	0.78	684
5	0.8	255	0	20	0.8	275
6	0.875	72	0	18	0.875	90
7	0.833	42	NaN	0	NaN	42
8	0.5	16	NaN	0	NaN	16

Khan, Sarah

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.739	1283	0	168	0.739	1451
2	0.739	1760	0	210	0.739	1970
3	0.745	1125	0	87	0.745	1212
4	0.833	628	0	36	0.833	664
5	0.809	325	0	10	0.809	335
6	0.767	60	NaN	0	NaN	60
7	0.69	42	NaN	0	NaN	42
9	1	27	NaN	0	NaN	27

Kirkland, Patricia Ann

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.743	1323	0	181	0.743	1504
2	0.765	1764	0	200	0.765	1964
3	0.758	1086	0	111	0.758	1197
4	0.782	696	0	32	0.782	728
5	0.765	230	0	5	0.765	235
6	0.857	84	0	6	0.857	90
7	0.8	35	NaN	0	NaN	35
8	1	8	NaN	0	NaN	8

Lazarev, Egor

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.721	1324	0	174	0.721	1498
2	0.768	1716	0	204	0.768	1920
3	0.734	1161	0	114	0.734	1275
4	0.766	620	0	28	0.766	648
5	0.663	255	0	15	0.663	270
6	0.845	84	0	6	0.845	90
7	0.8	35	NaN	0	NaN	35
8	1	16	NaN	0	NaN	16
9	0	9	NaN	0	NaN	9

Lozano, Andrea Patricia

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.737	1295	0	182	0.737	1477
2	0.735	1652	0	222	0.735	1874
3	0.798	1173	0	141	0.798	1314
4	0.772	628	0	28	0.772	656
5	0.754	305	NaN	0	NaN	305
6	0.889	90	NaN	0	NaN	90
7	0.8	35	NaN	0	NaN	35
10	1	10	NaN	0	NaN	10

Luby, Ryan Patrick

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.722	1279	0	189	0.722	1468
2	0.751	1790	0	180	0.751	1970
3	0.748	1200	0	93	0.748	1293
4	0.812	580	0	44	0.812	624
5	0.832	220	0	35	0.832	255
6	0.863	102	0	6	0.863	108
7	0.914	35	NaN	0	NaN	35
8	0	8	NaN	0	NaN	8

Marquez Pena, Javier

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.75	1343	0	184	0.75	1527
2	0.734	1670	0	214	0.734	1884
3	0.753	1221	0	78	0.753	1299
4	0.725	628	0	52	0.725	680
5	0.846	240	0	5	0.846	245
6	0.726	84	NaN	0	NaN	84
7	0.833	42	NaN	0	NaN	42

Moreno, Edgar Samuel

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.728	1299	0	175	0.728	1474
2	0.762	1700	0	206	0.762	1906
3	0.751	1188	0	90	0.751	1278
4	0.761	656	0	28	0.761	684
5	0.751	285	0	15	0.751	300
6	0.795	78	0	6	0.795	84
7	0.6	35	NaN	0	NaN	35

Pan, Yilin

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.743	1342	0	207	0.743	1549
2	0.75	1792	0	158	0.75	1950
3	0.762	1128	0	102	0.762	1230
4	0.778	612	0	32	0.778	644
5	0.715	235	0	30	0.715	265
6	0.725	102	NaN	0	NaN	102
7	1	21	NaN	0	NaN	21

Rink, Anselm Frieder

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.726	1350	0	182	0.726	1532
2	0.751	1664	0	208	0.751	1872
3	0.766	1152	0	99	0.766	1251
4	0.818	632	0	20	0.818	652
5	0.74	315	0	20	0.74	335
6	0.811	90	NaN	0	NaN	90
7	0	21	NaN	0	NaN	21
8	0.5	8	NaN	0	NaN	8

Sacramone-Lutz, Gabriella

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.743	1276	0	200	0.743	1476
2	0.734	1652	0	170	0.734	1822
3	0.753	1236	0	114	0.753	1350
4	0.769	648	0	24	0.769	672
5	0.746	280	0	15	0.746	295
6	0.781	96	0	6	0.781	102
7	0.75	28	0	7	0.75	35
9	1	9	NaN	0	NaN	9

Sharma, Kunaal

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.728	1287	0	196	0.728	1483
2	0.762	1678	0	186	0.762	1864
3	0.755	1152	0	129	0.755	1281
4	0.693	652	0	32	0.693	684
5	0.8	275	0	15	0.8	290
6	0.812	138	NaN	0	NaN	138
7	1	21	NaN	0	NaN	21

Snegovaya, Maria

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.751	1343	0	186	0.751	1529
2	0.748	1694	0	196	0.748	1890
3	0.754	1173	0	93	0.754	1266
4	0.747	660	0	48	0.747	708
5	0.829	210	0	15	0.829	225
6	0.643	84	0	6	0.643	90
7	0.571	28	NaN	0	NaN	28
8	1	16	NaN	0	NaN	16
9	1	9	NaN	0	NaN	9

Spry, Amber Denise

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.735	1334	0	183	0.735	1517
2	0.722	1710	0	212	0.722	1922
3	0.783	1176	0	105	0.783	1281
4	0.78	636	0	44	0.78	680
5	0.779	240	0	10	0.779	250
6	0.856	90	NaN	0	NaN	90
7	1	21	NaN	0	NaN	21

Tattersall, Laura

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.737	1320	0	166	0.737	1486
2	0.741	1730	0	200	0.741	1930
3	0.736	1137	0	132	0.736	1269
4	0.77	556	0	32	0.77	588
5	0.766	295	0	10	0.766	305
6	0.814	102	0	18	0.814	120
7	0.893	28	NaN	0	NaN	28
8	0.5	16	NaN	0	NaN	16
9	1	9	NaN	0	NaN	9
10	1	10	NaN	0	NaN	10

Warren, Shana

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.745	1330	0	178	0.745	1508
2	0.756	1656	0	182	0.756	1838
3	0.768	1200	0	141	0.768	1341
4	0.742	652	0	36	0.742	688
5	0.838	240	0	30	0.838	270
6	0.843	102	NaN	0	NaN	102
7	1	14	NaN	0	NaN	14

Zelizer, Adam Philip

Block j (# Persons)	$\mathbb{E}[d_i(Z=1)]$	$N_{j,z=1}$	$\mathbb{E}[d_i(Z=0)]$	$N_{j,z=0}$	$\mathbb{E}[d_i(1)] - \mathbb{E}[d_i(0)]$	N_j
1	0.737	1344	0	189	0.737	1533
2	0.754	1672	0	178	0.754	1850
3	0.765	1218	0	126	0.765	1344
4	0.777	588	0	60	0.777	648
5	0.733	225	0	15	0.733	240
6	0.9	90	0	6	0.9	96
7	0.905	42	NaN	0	NaN	42
8	1	8	NaN	0	NaN	8

Problem 2, Part F

Define the CACE in this context. Estimate the CACE and its 95% confidence interval. To what extent are these results changed when you control for covariates?

		No Covariate Adjustment						With Covariate Adjustment					
			RI p-values (for ITT)			Adj. 95% CI			RI p-values (for ITT)			Adj. 95% CI	
Student	k	$\widehat{CACE}$	Two-tailed	Greater	Lesser	LB	UB	$\widehat{CACE}$	Two-tailed	Greater	Lesser	LB	UB
Anderson-Hill, Shayna	0.16	0.00	0.00	1.00	-0.36	0.69	0.15	0.00	0.00	1.00	-0.37	0.68	
Ferrerosa-Young, Carolina	0.16	0.00	0.00	1.00	-0.36	0.68	0.15	0.00	0.00	1.00	-0.37	0.67	
Foos, Florian	0.12	0.00	0.00	1.00	-0.40	0.64	0.11	0.00	0.00	1.00	-0.41	0.64	
Khan, Sarah	0.14	0.00	0.00	1.00	-0.38	0.66	0.13	0.00	0.00	1.00	-0.39	0.65	
Kirkland, Patricia Ann	0.16	0.00	0.00	1.00	-0.37	0.68	0.15	0.00	0.00	1.00	-0.37	0.67	
Lazarev, Egor	0.14	0.00	0.00	1.00	-0.39	0.66	0.13	0.00	0.00	1.00	-0.40	0.65	
Lozano, Andrea Patricia	0.15	0.00	0.00	1.00	-0.37	0.68	0.15	0.00	0.00	1.00	-0.37	0.67	
Luby, Ryan Patrick	0.15	0.00	0.00	1.00	-0.37	0.67	0.14	0.00	0.00	1.00	-0.38	0.66	
Marquez Pena, Javier	0.14	0.00	0.00	1.00	-0.39	0.66	0.14	0.00	0.00	1.00	-0.39	0.66	
Moreno, Edgar Samuel	0.14	0.00	0.00	1.00	-0.39	0.66	0.13	0.00	0.00	1.00	-0.39	0.66	
Pan, Yilin	0.15	0.00	0.00	1.00	-0.37	0.67	0.15	0.00	0.00	1.00	-0.37	0.67	
Rink, Anselm Frieder	0.13	0.00	0.00	1.00	-0.40	0.65	0.12	0.00	0.00	1.00	-0.40	0.65	
Sacramone-Lutz, Gabriella	0.17	0.00	0.00	1.00	-0.36	0.69	0.16	0.00	0.00	1.00	-0.37	0.68	
Sharma, Kunaal	0.12	0.00	0.00	1.00	-0.40	0.64	0.11	0.00	0.00	1.00	-0.41	0.64	
Snegovaya, Maria	0.15	0.00	0.00	1.00	-0.37	0.67	0.14	0.00	0.00	1.00	-0.38	0.67	
Spry, Amber Denise	0.14	0.00	0.00	1.00	-0.39	0.66	0.13	0.00	0.00	1.00	-0.39	0.65	
Tattersall, Laura	0.15	0.00	0.00	1.00	-0.37	0.68	0.15	0.00	0.00	1.00	-0.37	0.67	
Warren, Shana	0.14	0.00	0.00	1.00	-0.39	0.66	0.13	0.00	0.00	1.00	-0.39	0.66	
Zelizer, Adam Philip	0.16	0.00	0.00	1.00	-0.36	0.68	0.15	0.00	0.00	1.00	-0.37	0.67	

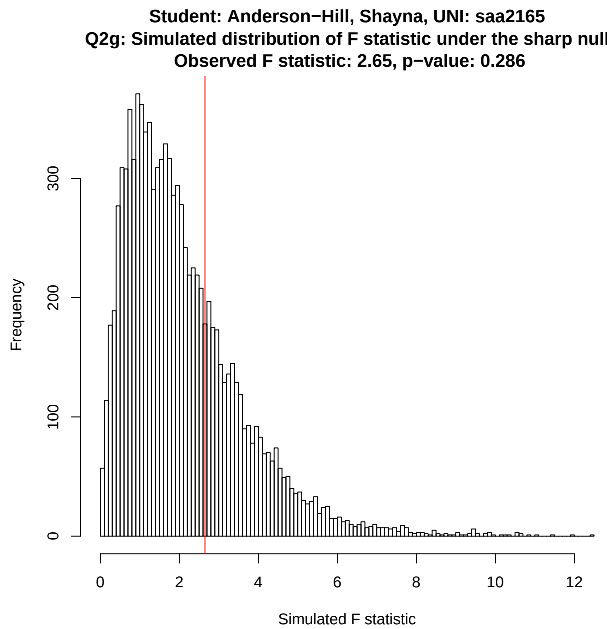
Problem 2, Part G

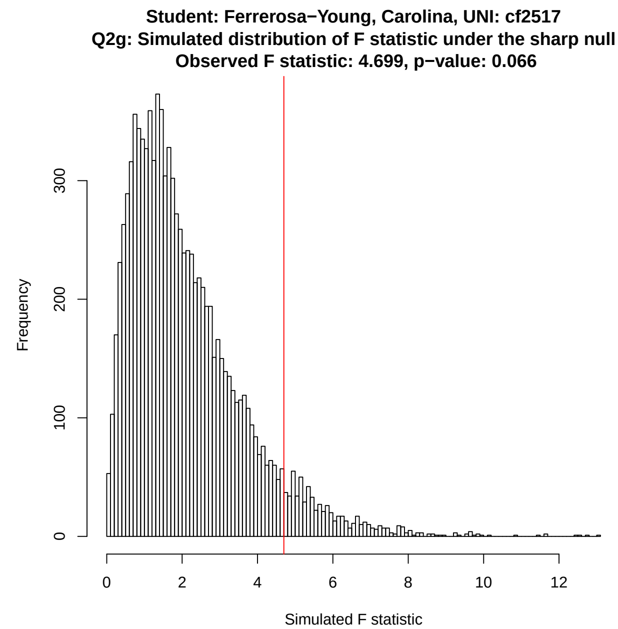
Conduct a randomization check in which treat2 is predicted by age, precinct, and voting in previous elections. Interpret the results.

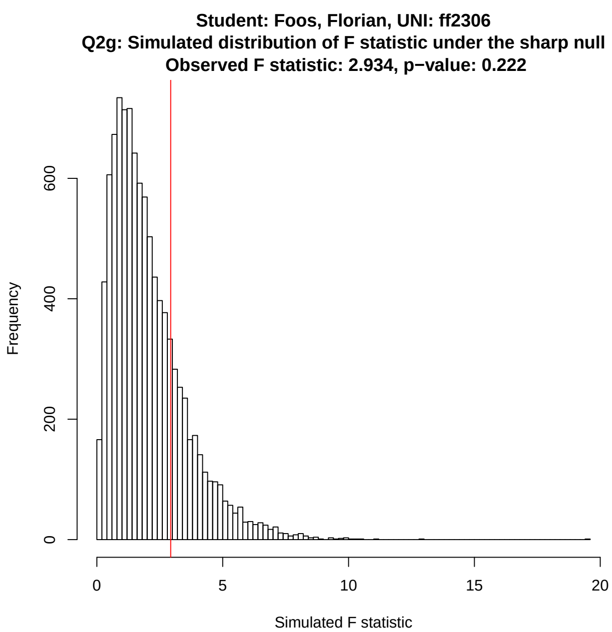
General Solution:

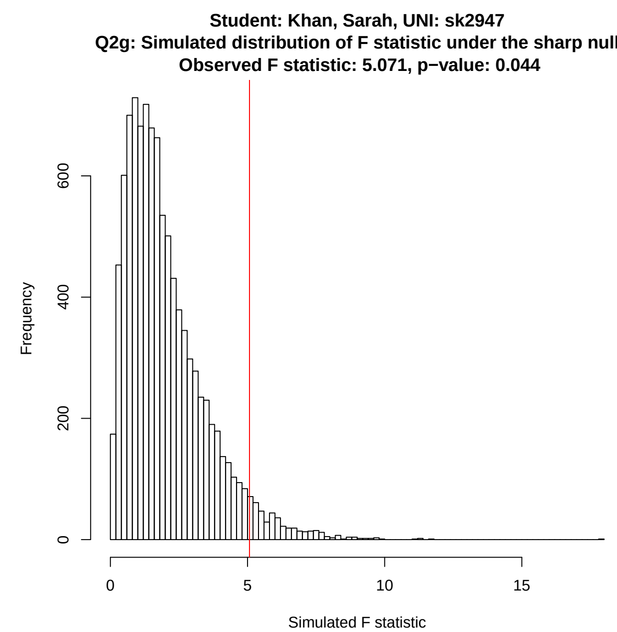
- Use randomization inference to test the null hypothesis that the covariates predict treatment assignment no better than would be expected by chance. The test statistic is the F statistic, which tells us the goodness of fit of a regression model (with all covariates) vs. a restricted model that contains only an intercept on the right hand side.
- First, regress treatment assignment (treat2) on four covariates (age, precinct, vote02, vote00); grab F statistic.
- Simulate permutations of treatment assignment, accounting for clustered random assignment (clustvar=hhid). For each simulated vector of treatment assignments, regress the simulated treatment assignment vector on the four covariates, and save the F statistic from every iteration.
- Compute a p-value: the probability of obtaining an F statistic under the null hypothesis at least as large as the one obtained from the actual experiment. If the p-value is small (e.g., p=0.01), then this tells us the imbalance is greater than one would expect by chance (FEDAI, pp. 107-08).

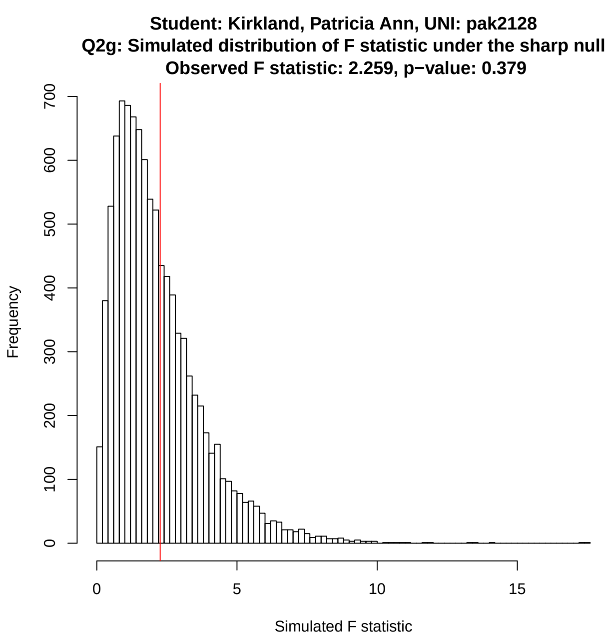
Student-Specific Solutions:

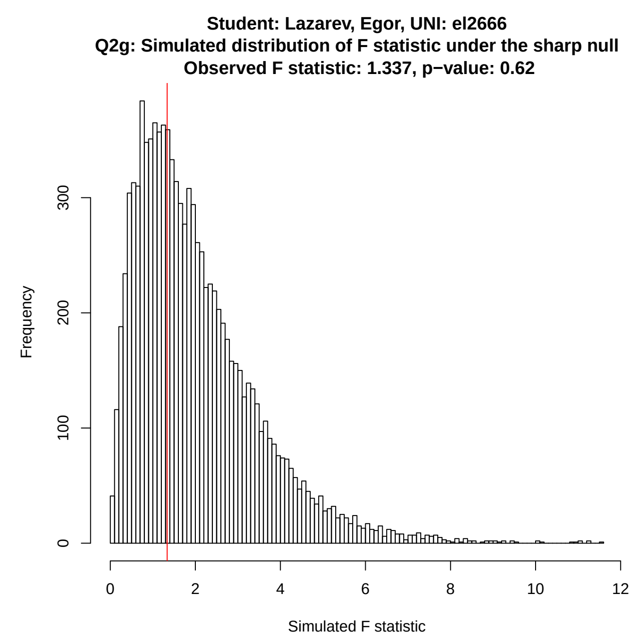

Student: Anderson−Hill, Shayna, UNI: saa2165
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.65, p−value: 0.286
Frequency
Simulated F statistic


Student: Ferrerosa−Young, Carolina, UNI: cf2517
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 4.699, p−value: 0.066
Frequency
Simulated F statistic

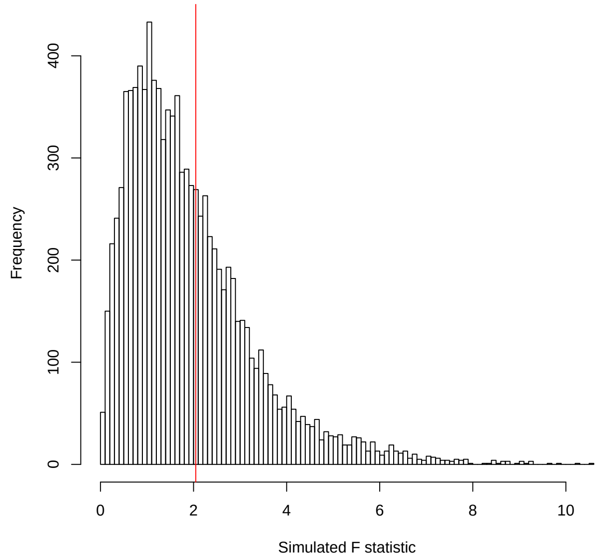

Student: Foos, Florian, UNI: ff2306
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.934, p−value: 0.222
Frequency
Simulated F statistic


Student: Khan, Sarah, UNI: sk2947
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 5.071, p−value: 0.044
Frequency
Simulated F statistic

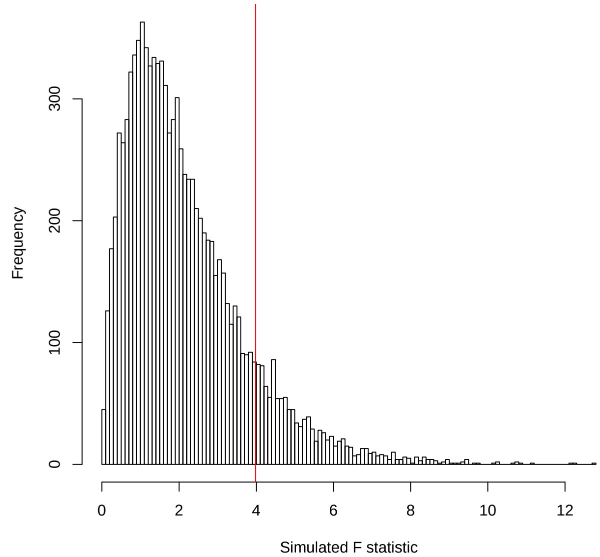

Student: Kirkland, Patricia Ann, UNI: pak2128
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.259, p−value: 0.379
Frequency
Simulated F statistic


Student: Lazarev, Egor, UNI: el2666
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 1.337, p−value: 0.62
Frequency
Simulated F statistic

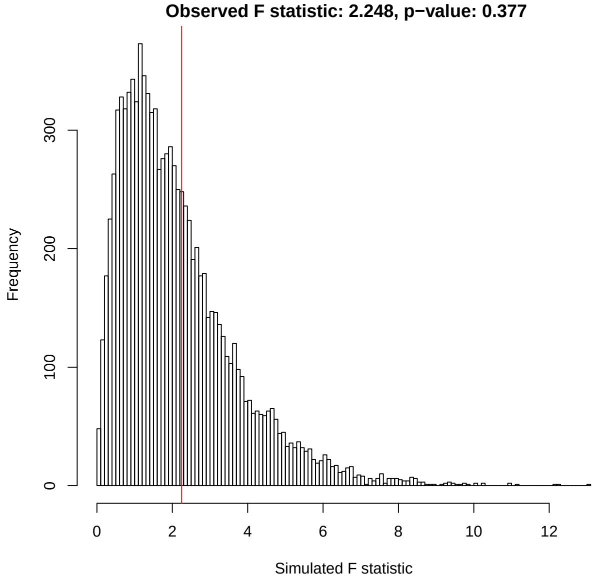
Student: Lozano, Andrea Patricia, UNI: apl2136
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.047, p−value: 0.37



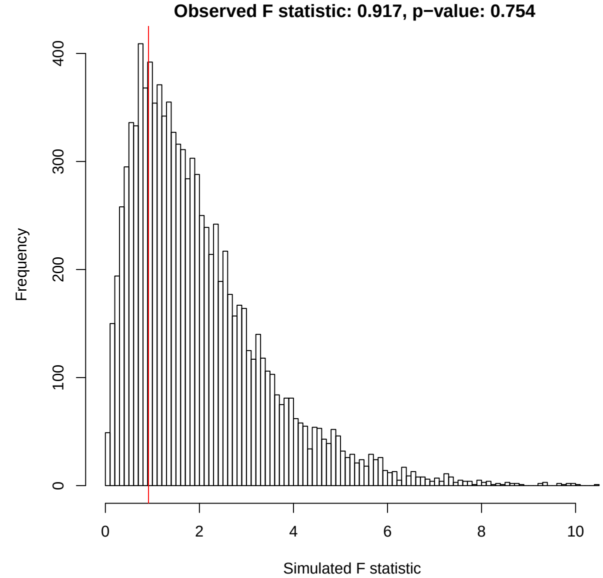
Student: Luby, Ryan Patrick, UNI: rpl2126
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 3.98, p−value: 0.118



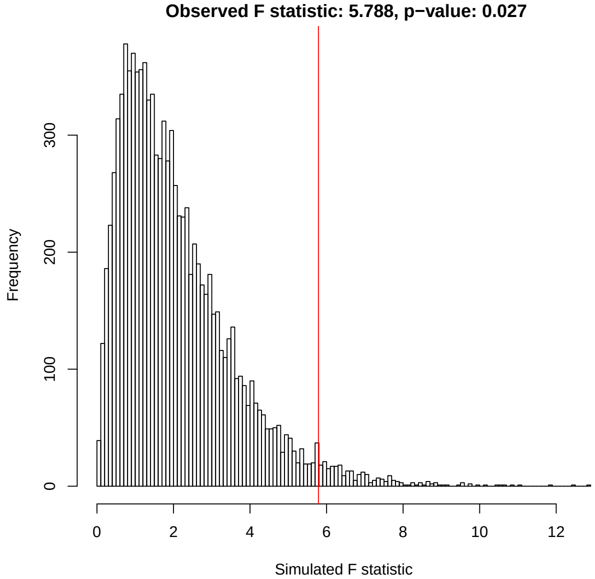
Student: Marquez Pena, Javier, UNI: jm3840
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.248, p−value: 0.377



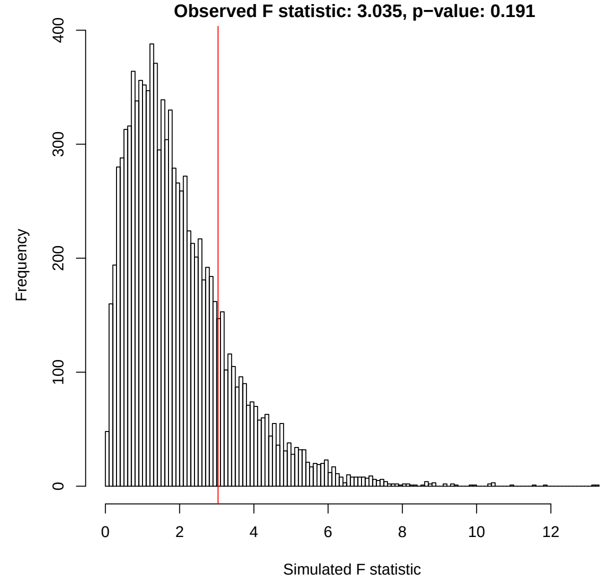
Student: Moreno, Edgar Samuel, UNI: esm2157
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 0.917, p−value: 0.754



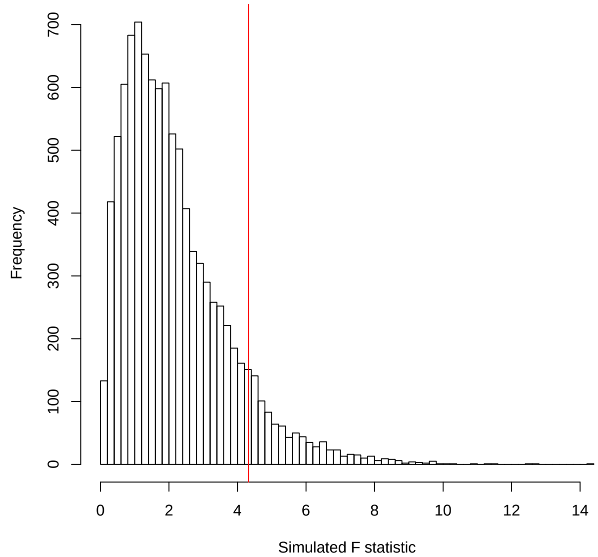
Student: Pan, Yilin, UNI: yp2266
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 5.788, p−value: 0.027



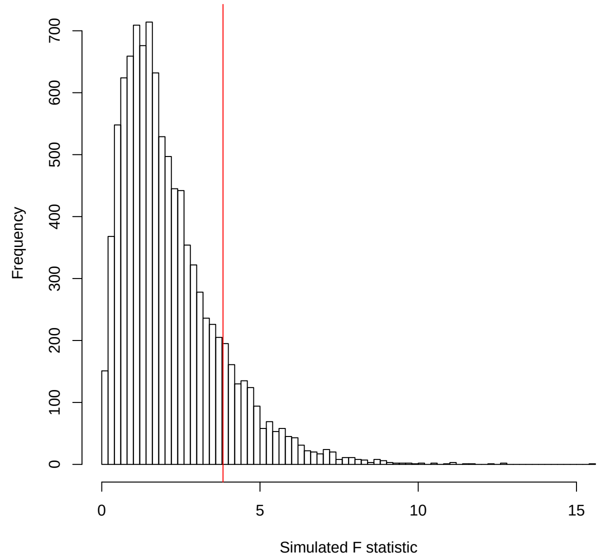
Student: Rink, Anselm Frieder, UNI: afr2132
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 3.035, p−value: 0.191



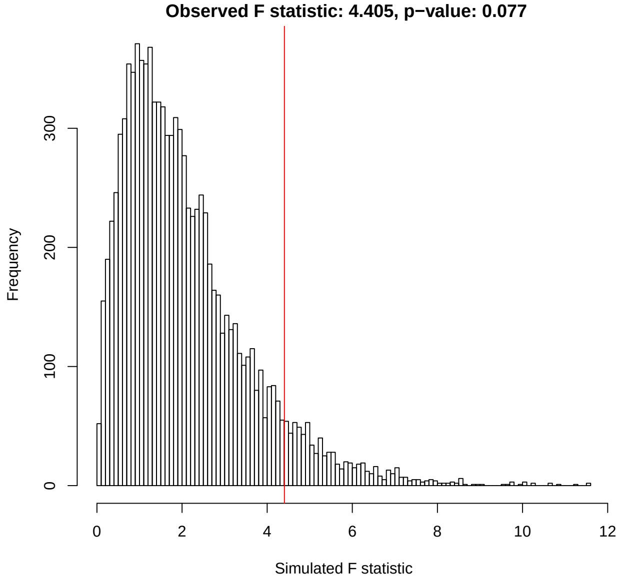
Student: Sacramone−Lutz, Gabriella, UNI: gs2580
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 4.319, p−value: 0.091



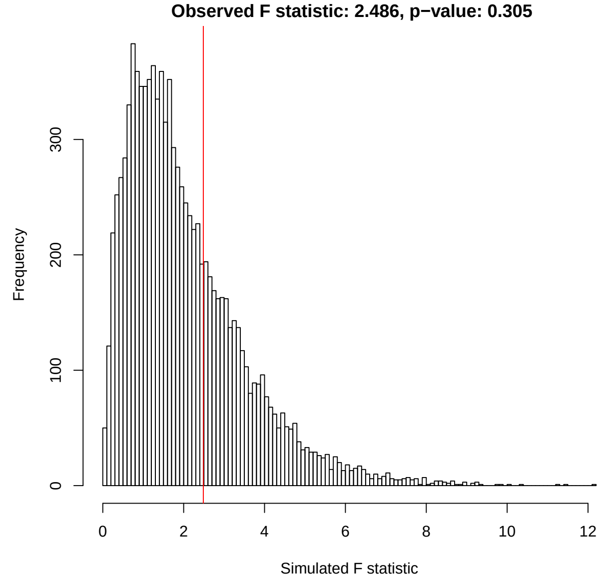
Student: Sharma, Kunaal, UNI: ks2481
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 3.832, p−value: 0.136



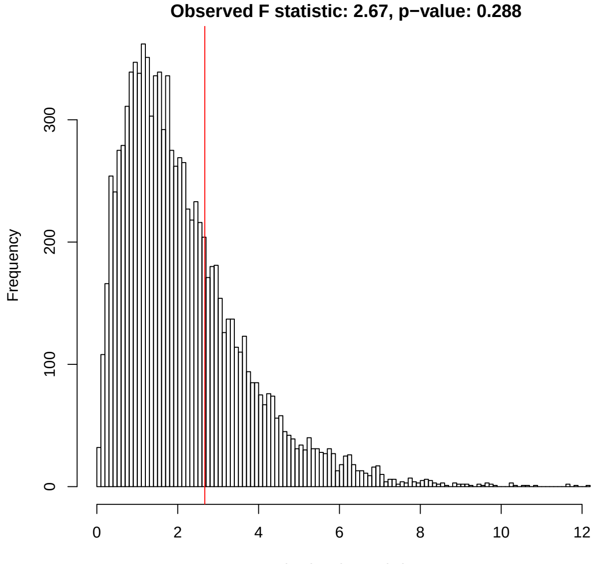
Student: Snegovaya, Maria, UNI: ms4391
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 4.405, p−value: 0.077



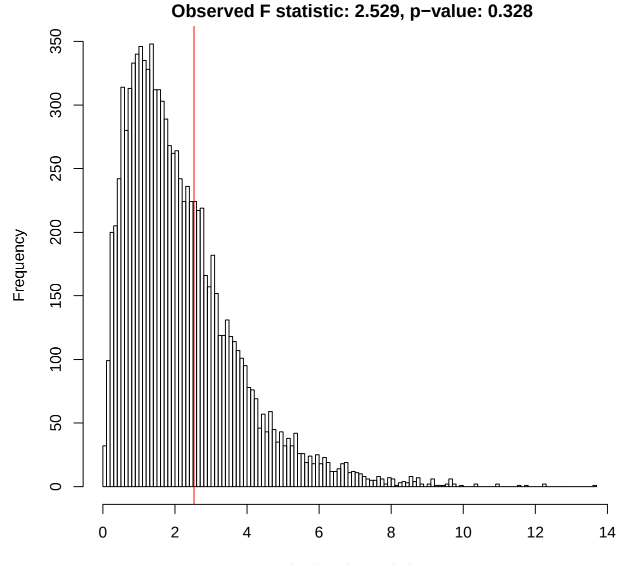
Student: Spry, Amber Denise, UNI: ads2183
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.486, p−value: 0.305

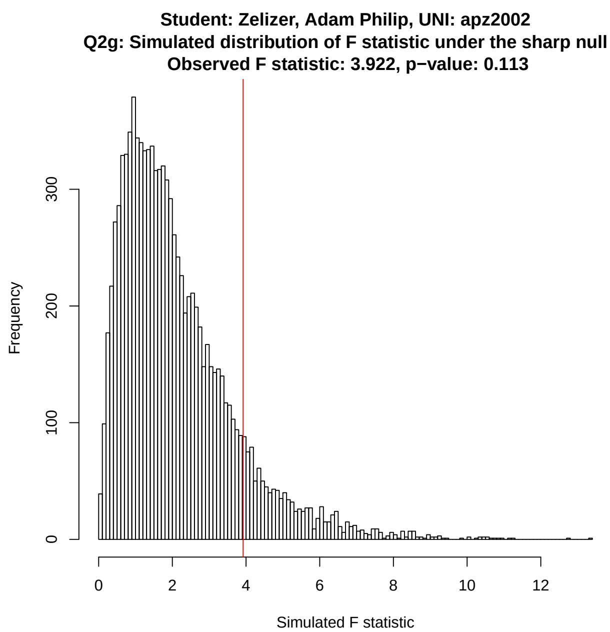


Student: Tattersall, Laura, UNI: lt2467
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.67, p−value: 0.288



Student: Warren, Shana, UNI: sw2647
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 2.529, p−value: 0.328




Student: Zelizer, Adam Philip, UNI: apz2002
Q2g: Simulated distribution of F statistic under the sharp null
Observed F statistic: 3.922, p−value: 0.113
Frequency
0
100
200
300
0
2
4
6
8
10
12
Simulated F statistic