Problem Set 3

Alex, Micah, Scott, and David

02/24/2025

Contents

1 Peruvian Recycling

1.1 Recycling bin ATE

In Column 3 of Table 4A, what is the estimated ATE of providing a recycling bin on the average weight of recyclables turned in per household per week, during the six-week treatment period? Provide a 95% confidence interval. Provide a short narrative using inline R code, such as r inline_reference.

```
# Use this code chunk to show your code work (if needed)
ate_bin <- 0.187  # Estimated ATE for recycling bin
se_bin <- 0.032   # Standard error

ci_bin <- ate_bin + c(-1.96, 1.96) * se_bin
ci_bin
```

```
## [1] 0.12428 0.24972
```

Answer: The estimated ATE of providing a recycling bin on the average weight of recyclables turned in per household per week is 0.187, with a 95% confidence interval of (0.12428,0.24972).

1.2 SMS ATE

In Column 3 of Table 4A, what is the estimated ATE of sending a text message reminder on the average weight of recyclables turned in per household per week? Provide a 95% confidence interval and provide a short narrative using inline R code.

```
# Use this code chunk to show your code work (if needed)
ate_sms <- -0.024  # Estimated ATE for SMS
se_sms <- 0.039    # Standard error

ci_sms <- ate_sms + c(-1.96, 1.96) * se_sms
ci_sms
```

```
## [1] -0.10044  0.05244
```

Answer: The estimated ATE of sending an SMS reminder is -0.024, with a 95% confidence interval of (-0.10044,0.05244).

1.3 What outcomes does a recycling bin affect?

Which outcome measures in Table 4A show statistically significant effects (at the 5% level) of providing a recycling bin? How are you dealing with the issue that there are several different tests that have been run, and that you are reading? How, if at all, do the authors deal with this?

Answer: The recycling bin has statistically significant effects on: - The number of bins turned in per week (0.115) - The weight in kg of recyclables turned in per week (0.187) - The market value of recyclables given per week (0.108)

The issue of multiple tests is addressed by considering significance levels and reporting p-values. The authors do not explicitly mention adjusting for multiple comparisons.

1.4 What outcomes does a SMS affect?

Which outcome measures in Table 4A show statistically significant effects (at the 5% level) of sending text messages? Now that you have read across two different treatments, and many outcomes, what, if anything do the p-values mean to you? Does this feel like p-hacking or doing careful investigation?

Answer: The SMS treatment does not have statistically significant effects on any outcome at the 5% level. Given multiple tests, the p-values should be interpreted cautiously. This does not feel like p-hacking because the study pre-specifies outcomes and treatment variables.

1.5 Marginal effects

Suppose that, during the two weeks before treatment, household A turns in 2kg per week more recyclables than household B does, and suppose that both households are otherwise identical (including being in the same treatment group). From the model, how much more recycling do we predict household A to have than household B, per week, during the six weeks of treatment? Provide only a point estimate, as the confidence interval would be a bit complicated. This question is designed to test your understanding of slope coefficients in regression.

```r
# Use this code chunk to show your code work (if needed)
marginal_effect <- 2 * 0.281
marginal_effect
```

```
## [1] 0.562
```

Answer: The predicted additional recycling for Household A compared to Household B per week during treatment is 0.562 kg.

1.6 Covariates or confounders?

Suppose that the variable "percentage of visits turned in bag, baseline" had been left out of the regression reported in Column 1. What would you expect to happen to the results on providing a recycling bin? Would you expect an increase or decrease in the estimated ATE? Would you expect an increase or decrease in the standard error? Explain your reasoning.

```r
# Use this code chunk to show your code work (if needed)
```

Answer: If "percentage of visits turned in bag, baseline" were omitted, the estimated ATE for the recycling bin would likely increase due to omitted variable bias. The standard error would likely increase because a relevant control variable would be missing, leading to higher variance in estimates.

1.7 Bad control or useful subset?

In column 1 of Table 4A, would you say the variable "has cell phone" is a bad control? Explain your reasoning, and engage both with the definition of a bad control, and also the implications of including a bad control in a model.

Answer: "has cell phone" may be a bad control if it is endogenous to the recycling decision. Including it could introduce bias rather than properly controlling for pre-existing differences.

1.8 What happens if you remove "has cell phone"?

If we were to remove the "has cell phone" variable from the regression, what would you expect to happen to the coefficient on "Any SMS message"? Would it go up or down? Explain your reasoning.

```r
# Use this code chunk to show your code work (if needed)
```

Answer: If we remove "has cell phone" from the regression, the coefficient on "Any SMS message" would likely increase because cell phone ownership is likely positively correlated with responsiveness to SMS messages.

2 Multifactor Experiments

2.1 Experiment design?

What is the full experimental design for this experiment? Tell us the dimensions, such as 2x2x3. The full results appear in Panel 4B. We'll note that the dimensions of an experiment are defined in terms of the *treatments that the experiment assigns*, not in terms of other features about the data.

Answer: The experiment follows a 3x3 factorial design, where there are two independent treatment factors:

1. Recycling bin treatment (3 levels):

- No bin
- Bin with a sticker
- Bin without a sticker

2. SMS message treatment (3 levels):

- No SMS message
- Personal SMS message
- Generic SMS message

This results in a total of 9 experimental groups (3x3).

2.2 Baseline for interpretation

In the results of Table 4B, describe the baseline category. That is, in English, how would you describe the attributes of the group of people for whom all dummy variables are equal to zero?

Answer: The baseline category in Table 4B consists of households that: - Did not receive a recycling bin - Did not receive any SMS message

In regression models, this means that all dummy variables for treatment groups are equal to zero, and the estimated effects for other treatment groups are interpreted relative to this baseline.

2.3 Bin without sticker effect

In column (1) of Table 4B, interpret the magnitude of the coefficient on "bin without sticker." What does it mean?

Answer: In Column (1) of Table 4B, the coefficient for "Bin without sticker" is 0.035. This means that receiving a recycling bin without a sticker increased the percentage of visits where households turned in a bag by 3.5 percentage points, compared to the baseline group (households without a bin).

2.4 With or without a sticker?

In column (1) of Table 4B, which seems to have a stronger treatment effect, the recycling bin with message sticker, or the recycling bin without sticker? How large is the magnitude of the estimated difference?

Answer: - The coefficient for Bin with sticker is 0.055 - The coefficient for Bin without sticker is 0.035 - The difference in effect size is 0.055 - 0.035 = 0.020

Thus, the bin with a sticker had a stronger effect by about 2 percentage points in increasing the percentage of visits where a bag was turned in.

2.5 Statistical significantly different with or without a sticker?

Is this difference you just described statistically significant? Explain which piece of information in the table allows you to answer this question.

Answer: - The Bin with sticker effect (0.055) is statistically significant at the 1% level - The Bin without sticker effect (0.035) is statistically significant at the 5% level - To determine whether the difference (2 percentage points) is statistically significant, we would need a formal hypothesis test

Since the table does not directly provide the p-value for the difference, we cannot confirm statistical significance solely from Table 4B, but we can infer that it is likely meaningful.

2.6 Fully saturated?

Notice that Table 4C is described as results from "fully saturated" models. What does this mean? What does David Reiley propose this definition means to him in the async lecture? What do the authors seem to think it means to them? Looking at the list of variables in the table, explain in what sense the model is "saturated."

Answer: A fully saturated model includes all possible interaction effects among the treatment groups.

In Table 4C, each possible combination of bin type and SMS message has its own coefficient. This ensures that treatment effects are estimated separately for each combination instead of assuming additive effects.

Interpretation by Different Sources David Reiley's definition: A fully saturated model means that each treatment combination has its own separate coefficient rather than assuming treatment effects add up independently. Authors' definition: The model allows all interactions to be explicitly estimated, meaning it does not impose any assumptions about how the treatments combine.

Thus, the model is saturated because it fully accounts for all treatment combinations, leaving no room for further interaction terms or omitted variable concerns.

3 Now! Do it with data

3.1 Treatment only model

A. For simplicity, let's start by measuring the effect of providing a recycling bin, ignoring the SMS message treatment (and ignoring whether there was a sticker on the bin or not). Run a regression of Y on only the bin treatment dummy, so you estimate a simple difference in means. Provide a 95% confidence interval for the treatment effect, using **of course** robust standard errors (use these throughout) and provide a brief narrative using R inline statements.

```
mod_basic <- lm(avg_bins_treat ~ bin, data = d_clean)
coeftest(mod_basic, vcov = vcovHC(mod_basic, type = "HC1"))
```

```
##
## t test of coefficients:
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 0.635888   0.011471 55.4367 < 2.2e-16 ***
## bin         0.133316   0.020673  6.4488 1.448e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
confint(mod_basic)
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.61285044 0.6589249
## bin         0.09364785 0.1729850
```

```
## TODO, include code to print these estimates and robust standard errors to the screen
## remove these comments once you've written that code.
```

Narrative: The effect of providing a recycling bin is estimated using a simple linear regression model. The 95% confidence interval for the treatment effect is provided using robust standard errors.

3.2 Treatment and pre-treatment values

Now add the pre-treatment value of Y as a covariate. Provide a 95% confidence interval for the treatment effect. Explain how and why this confidence interval differs from the previous one.

```
mod_pretreat <- lm(avg_bins_treat ~ bin + base_avg_bins_treat, data = d_clean)
coeftest(mod_pretreat, vcov = vcovHC(mod_pretreat, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                     Estimate Std. Error t value  Pr(>|t|)
## (Intercept)         0.352055   0.020711 16.9987 < 2.2e-16 ***
## bin                 0.124242   0.017119  7.2576  5.86e-13 ***
## base_avg_bins_treat 0.390001   0.029512 13.2152 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
confint(mod_pretreat)
```

```
##                          2.5 %    97.5 %
## (Intercept)         0.32507625 0.3790345
## bin                 0.09155241 0.1569323
```

```
## base_avg_bins_treat 0.36365684 0.4163454
```

Answer: Adding the pre-treatment values as a covariate reduces variance in the model, leading to a narrower confidence interval for the treatment effect.

3.3 Add state fixed effects

Now add the street fixed effects. (You'll need to use the R command `factor()` You can do this either within the `lm` call, or you can move this factoring up in the data pipeline so that it persists through the rest of your analysis. The only thing we would recommend that you *not* do is to engineer a new, persistent feature at this point.) Provide a 95% confidence interval for the treatment effect and provide a brief narrative using r inline statements.

```
mod_fixed_effects <- lm(avg_bins_treat ~ bin + base_avg_bins_treat + factor(street), data = d_clean)
coeftest(mod_fixed_effects, vcov = vcovHC(mod_fixed_effects, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                       Estimate Std. Error t value  Pr(>|t|)
## (Intercept)          0.3677440  0.0345745 10.6363 < 2.2e-16 ***
## bin                  0.1138868  0.0175138  6.5027 1.052e-10 ***
## base_avg_bins_treat  0.3737068  0.0269012 13.8918 < 2.2e-16 ***
## factor(street)2     -0.0946095  0.0581010 -1.6284 0.1036447
## factor(street)3     -0.0447175  0.1017270 -0.4396 0.6602984
## factor(street)4     -0.1350441  0.0628410 -2.1490 0.0317855 *
## factor(street)5     -0.0942115  0.0982573 -0.9588 0.3377923
## factor(street)6     -0.0507716  0.1058096 -0.4798 0.6314071
## factor(street)7     -0.0546521  0.1325565 -0.4123 0.6801801
## factor(street)8     -0.0540184  0.0562084 -0.9610 0.3366782
## factor(street)9     -0.0868541  0.0989790 -0.8775 0.3803469
## factor(street)10    -0.0275578  0.0816097 -0.3377 0.7356500
## factor(street)11    -0.0709719  0.1118462 -0.6345 0.5258131
## factor(street)15    -0.0200592  0.1071294 -0.1872 0.8514943
## factor(street)17     0.0396003  0.0872336  0.4540 0.6499210
## factor(street)20    -0.2359875  0.0728101 -3.2411 0.0012150 **
## factor(street)21    -0.0759775  0.0799868 -0.9499 0.3423192
## factor(street)22    -0.1252626  0.0685151 -1.8282 0.0676984 .
## factor(street)23    -0.1601861  0.0559034 -2.8654 0.0042191 **
## factor(street)26    -0.0277308  0.0670593 -0.4135 0.6792762
## factor(street)32     0.1560981  0.1795052  0.8696 0.3846484
## factor(street)37     0.1730332  0.1755399  0.9857 0.3244195
## factor(street)38    -0.0711148  0.0814907 -0.8727 0.3829720
## factor(street)40    -0.0623286  0.2308444 -0.2700 0.7871929
## factor(street)41     0.0127359  0.1578288  0.0807 0.9356953
## factor(street)42    -0.1448419  0.1175334 -1.2323 0.2180006
## factor(street)43     0.0531346  0.1298502  0.4092 0.6824486
## factor(street)44     0.1346742  0.1094424  1.2305 0.2186728
## factor(street)45    -0.0388791  0.1012805 -0.3839 0.7011218
## factor(street)46     0.0101859  0.0898380  0.1134 0.9097429
## factor(street)47     0.4194175  0.2552953  1.6429 0.1006059
## factor(street)53     0.0362359  0.1467310  0.2470 0.8049750
## factor(street)58    -0.0305051  0.0641873 -0.4753 0.6346723
## factor(street)60     0.0280276  0.0681108  0.4115 0.6807615
## factor(street)61    -0.0085464  0.1410588 -0.0606 0.9516953
```

```
## factor(street)62      0.0574930  0.1216858  0.4725 0.6366551
## factor(street)63     -0.0284624  0.0850189 -0.3348 0.7378369
## factor(street)64      0.0509659  0.0895826  0.5689 0.5694861
## factor(street)66     -0.0602193  0.0554458 -1.0861 0.2776011
## factor(street)67     -0.0129774  0.1422173 -0.0913 0.9273047
## factor(street)68      0.0727117  0.1111695  0.6541 0.5131661
## factor(street)69      0.0780759  0.0803588  0.9716 0.3314012
## factor(street)70     -0.0085089  0.0795166 -0.1070 0.9147964
## factor(street)72     -0.0472835  0.0813183 -0.5815 0.5610108
## factor(street)73     -0.1711941  0.0589375 -2.9047 0.0037269 **
## factor(street)74     -0.0650898  0.0702932 -0.9260 0.3545988
## factor(street)75     -0.0563252  0.0519035 -1.0852 0.2780006
## factor(street)77     -0.0332737  0.1599142 -0.2081 0.8351993
## factor(street)78     -0.0517554  0.0800982 -0.6461 0.5182756
## factor(street)79      0.3067109  0.4148285  0.7394 0.4597921
## factor(street)80      0.0116236  0.0514772  0.2258 0.8213858
## factor(street)81     -0.0310541  0.1097868 -0.2829 0.7773220
## factor(street)82      0.0327803  0.1551712  0.2113 0.8327172
## factor(street)83     -0.1188480  0.1146195 -1.0369 0.2999433
## factor(street)84     -0.0178708  0.0628131 -0.2845 0.7760587
## factor(street)85      0.0650530  0.0744167  0.8742 0.3821560
## factor(street)86     -0.0670392  0.0742080 -0.9034 0.3664515
## factor(street)88     -0.0975655  0.1024263 -0.9525 0.3409657
## factor(street)89     -0.0035891  0.0984552 -0.0365 0.9709248
## factor(street)91     -0.2108308  0.1277210 -1.6507 0.0989934 .
## factor(street)93     -0.0701460  0.0814932 -0.8608 0.3894999
## factor(street)94     -0.0370345  0.0796883 -0.4647 0.6421797
## factor(street)96     -0.0389422  0.1563872 -0.2490 0.8033841
## factor(street)98      0.0523747  0.0838543  0.6246 0.5323276
## factor(street)99      0.2314373  0.0930778  2.4865 0.0130022 *
## factor(street)100     0.0743365  0.1418641  0.5240 0.6003525
## factor(street)101     0.1972951  0.1133931  1.7399 0.0820649 .
## factor(street)102    -0.0853576  0.1430253 -0.5968 0.5507251
## factor(street)103     0.0861700  0.1364732  0.6314 0.5278654
## factor(street)105     0.1160225  0.0790947  1.4669 0.1426052
## factor(street)106     0.4004565  0.1289437  3.1057 0.0019318 **
## factor(street)107     0.0643492  0.1527775  0.4212 0.6736690
## factor(street)109    -0.0145341  0.1146899 -0.1267 0.8991738
## factor(street)110     0.1407639  0.1877299  0.7498 0.4534725
## factor(street)111     0.0013745  0.1486143  0.0092 0.9926218
## factor(street)112     0.1847875  0.0849550  2.1751 0.0297670 *
## factor(street)113     0.0817787  0.1613810  0.5067 0.6124051
## factor(street)115     0.2923848  0.1316328  2.2212 0.0264759 *
## factor(street)117     0.5573269  0.4039905  1.3796 0.1679167
## factor(street)118     0.0328714  0.2601589  0.1264 0.8994696
## factor(street)119    -0.0109633  0.1086446 -0.1009 0.9196347
## factor(street)120    -0.0585693  0.0888945 -0.6589 0.5100784
## factor(street)121     0.1781558  0.1628523  1.0940 0.2741320
## factor(street)122    -0.1607091  0.1118776 -1.4365 0.1510635
## factor(street)124     0.0394846  0.1071357  0.3685 0.7125137
## factor(street)125     0.2995963  0.1425814  2.1012 0.0357769 *
## factor(street)126    -0.1035186  0.1253187 -0.8260 0.4089030
## factor(street)127     0.0735025  0.3373537  0.2179 0.8275507
## factor(street)128     0.1453292  0.1114527  1.3040 0.1924367
```

```
## factor(street)129    0.3770692  0.4566026  0.8258 0.4090323
## factor(street)130    0.1760675  0.1328629  1.3252 0.1853001
## factor(street)131    0.2756383  0.1488438  1.8519 0.0642297 .
## factor(street)132    0.2346483  0.1001776  2.3423 0.0192861 *
## factor(street)133    0.3158659  0.1070747  2.9500 0.0032243 **
## factor(street)134   -0.0104975  0.1272001 -0.0825 0.9342376
## factor(street)136    0.1808409  0.1395728  1.2957 0.1952746
## factor(street)137    0.2550936  0.1125404  2.2667 0.0235425 *
## factor(street)138    0.1211170  0.1639105  0.7389 0.4600629
## factor(street)147    0.1927181  0.0651251  2.9592 0.0031297 **
## factor(street)148    0.1060159  0.2234346  0.4745 0.6352204
## factor(street)149    0.2817032  0.1163635  2.4209 0.0155933 *
## factor(street)151   -0.2634108  0.1623033 -1.6230 0.1047961
## factor(street)152    0.1246159  0.2398345  0.5196 0.6034203
## factor(street)153   -0.0277308  0.1101538 -0.2517 0.8012695
## factor(street)154   -0.0593654  0.1098657 -0.5403 0.5890340
## factor(street)155   -0.2449064  0.1743298 -1.4048 0.1602614
## factor(street)156    0.1957426  0.1769510  1.1062 0.2688079
## factor(street)157    0.1488075  0.0709921  2.0961 0.0362288 *
## factor(street)158    0.3980332  0.1998794  1.9914 0.0466103 *
## factor(street)160   -0.1214908  0.0823994 -1.4744 0.1405672
## factor(street)163   -0.0780089  0.0811343 -0.9615 0.3364567
## factor(street)164   -0.0978748  0.0561571 -1.7429 0.0815478 .
## factor(street)165    0.0096059  0.1151824  0.0834 0.9335463
## factor(street)166   -0.1476841  0.1055813 -1.3988 0.1620753
## factor(street)168    0.1509603  0.3754429  0.4021 0.6876746
## factor(street)170   -0.0074312  0.0687287 -0.1081 0.9139108
## factor(street)171   -0.0470975  0.0989437 -0.4760 0.6341374
## factor(street)172   -0.0253266  0.0467439 -0.5418 0.5880205
## factor(street)175    0.2512113  0.1656083  1.5169 0.1294894
## factor(street)179    0.0036892  0.1407129  0.0262 0.9790870
## factor(street)180   -0.0975873  0.0667055 -1.4630 0.1436755
## factor(street)182   -0.1455463  0.1116001 -1.3042 0.1923607
## factor(street)183   -0.0830115  0.0769900 -1.0782 0.2811023
## factor(street)185   -0.0620430  0.0912027 -0.6803 0.4964281
## factor(street)186    0.0035617  0.0893459  0.0399 0.9682066
## factor(street)187   -0.0098479  0.0756260 -0.1302 0.8964100
## factor(street)188   -0.0285789  0.0518649 -0.5510 0.5816924
## factor(street)189   -0.0551111  0.0632856 -0.8708 0.3839771
## factor(street)190    0.0111908  0.1151649  0.0972 0.9226016
## factor(street)191   -0.0155078  0.0909272 -0.1706 0.8645978
## factor(street)192   -0.1048446  0.0662945 -1.5815 0.1139617
## factor(street)193    0.0619534  0.1063710  0.5824 0.5603609
## factor(street)196   -0.2274375  0.1137584 -1.9993 0.0457444 *
## factor(street)197    0.0711737  0.3362765  0.2117 0.8324054
## factor(street)198    0.0834945  0.1080690  0.7726 0.4398708
## factor(street)200   -0.0442727  0.0604657 -0.7322 0.4641564
## factor(street)202    0.0067345  0.1272118  0.0529 0.9577867
## factor(street)203   -0.2260193  0.0798015 -2.8323 0.0046798 **
## factor(street)206   -0.2377688  0.0898236 -2.6471 0.0081991 **
## factor(street)207   -0.2367958  0.0448675 -5.2777 1.488e-07 ***
## factor(street)208   -0.1169355  0.1058541 -1.1047 0.2694622
## factor(street)209   -0.3096190  0.0864653 -3.5808 0.0003527 ***
## factor(street)210   -0.2296541  0.0822192 -2.7932 0.0052812 **
```

```
## factor(street)213    0.0423786  0.0808600  0.5241 0.6002823
## factor(street)215   -0.0589301  0.0733195 -0.8037 0.4216641
## factor(street)216   -0.4564162  0.3450323 -1.3228 0.1860837
## factor(street)217   -0.1794699  0.1154332 -1.5548 0.1202032
## factor(street)220   -0.2052784  0.1408042 -1.4579 0.1450644
## factor(street)221   -0.1506315  0.0967153 -1.5575 0.1195558
## factor(street)222   -0.1743467  0.1198281 -1.4550 0.1458726
## factor(street)223   -0.1795225  0.1149222 -1.5621 0.1184570
## factor(street)225   -0.0902501  0.1164657 -0.7749 0.4385092
## factor(street)227   -0.0466068  0.1525403 -0.3055 0.7599964
## factor(street)228    0.0496937  0.1822067  0.2727 0.7850941
## factor(street)229    0.0365969  0.1474181  0.2483 0.8039711
## factor(street)230    0.0100587  0.0898262  0.1120 0.9108534
## factor(street)232    0.0777992  0.2486789  0.3128 0.7544354
## factor(street)233    0.0524326  0.1319866  0.3973 0.6912310
## factor(street)235   -0.1110214  0.1548147 -0.7171 0.4734020
## factor(street)236   -0.0134474  0.0722321 -0.1862 0.8523353
## factor(street)238   -0.1081611  0.0905453 -1.1946 0.2324390
## factor(street)240   -0.2970694  0.0656420 -4.5256 6.467e-06 ***
## factor(street)241   -0.1036475  0.1794409 -0.5776 0.5636064
## factor(street)242    0.0396003  0.0583041  0.6792 0.4971074
## factor(street)243    0.1582858  0.2612410  0.6059 0.5446674
## factor(street)244    0.0907676  0.1099975  0.8252 0.4093931
## factor(street)246   -0.1372508  0.0942168 -1.4568 0.1453801
## factor(street)247   -0.0028026  0.0550439 -0.0509 0.9593993
## factor(street)248   -0.0056622  0.0767263 -0.0738 0.9411805
## factor(street)249    0.0867767  0.1258116  0.6897 0.4904604
## factor(street)250    0.0454185  0.0507941  0.8942 0.3713655
## factor(street)253    0.1090277  0.0859432  1.2686 0.2047680
## factor(street)254    0.1011168  0.1161127  0.8709 0.3839665
## factor(street)255    0.0275976  0.1032349  0.2673 0.7892507
## factor(street)256    0.0247982  0.0632068  0.3923 0.6948637
## factor(street)257   -0.1036156  0.0713419 -1.4524 0.1465916
## factor(street)258    0.0966864  0.0575984  1.6786 0.0934193 .
## factor(street)259    0.7015603  0.2371876  2.9578 0.0031436 **
## factor(street)260   -0.0321524  0.0757366 -0.4245 0.6712372
## factor(street)261   -0.0435433  0.0749100 -0.5813 0.5611369
## factor(street)262    0.0356448  0.1052055  0.3388 0.7347968
## factor(street)263    0.0268892  0.0967800  0.2778 0.7811723
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
confint(mod_fixed_effects)
```

```
##                           2.5 %      97.5 %
## (Intercept)          0.305731652 0.429756389
## bin                  0.080428729 0.147344834
## base_avg_bins_treat  0.345602946 0.401810558
## factor(street)2     -0.294622280 0.105403181
## factor(street)3     -0.253677536 0.164242624
## factor(street)4     -0.354615547 0.084527394
## factor(street)5     -0.313570512 0.125147513
## factor(street)6     -0.282678090 0.181134835
## factor(street)7     -0.344529867 0.235225713
## factor(street)8     -0.233337002 0.125300257
```

```
## factor(street)9      -0.279042647  0.105334450
## factor(street)10     -0.184285613  0.129170022
## factor(street)11     -0.263387779  0.121443930
## factor(street)15     -0.343089357  0.302971033
## factor(street)17     -0.225913362  0.305114027
## factor(street)20     -0.467905870 -0.004069081
## factor(street)21     -0.341708173  0.189753271
## factor(street)22     -0.260477411  0.009952170
## factor(street)23     -0.302100194 -0.018271923
## factor(street)26     -0.398755846  0.343294238
## factor(street)32     -0.215149901  0.527346045
## factor(street)37     -0.116913969  0.462980363
## factor(street)38     -0.360991470  0.218761928
## factor(street)40     -0.433580101  0.308922840
## factor(street)41     -0.276989243  0.302460974
## factor(street)42     -0.516428639  0.226744868
## factor(street)43     -0.236603226  0.342872326
## factor(street)44     -0.097260978  0.366609351
## factor(street)45     -0.361553370  0.283795203
## factor(street)46     -0.255328696  0.275700495
## factor(street)47      0.096379642  0.742455312
## factor(street)53     -0.334737608  0.407209396
## factor(street)58     -0.215892158  0.154881863
## factor(street)60     -0.140987331  0.197042455
## factor(street)61     -0.380380771  0.363287956
## factor(street)62     -0.189369117  0.304355129
## factor(street)63     -0.260352276  0.203427444
## factor(street)64     -0.214540370  0.316472129
## factor(street)66     -0.307014502  0.186575843
## factor(street)67     -0.232319772  0.206364878
## factor(street)68     -0.159041423  0.304464891
## factor(street)69     -0.122128363  0.278280081
## factor(street)70     -0.187854223  0.170836487
## factor(street)72     -0.247298407  0.152731396
## factor(street)73     -0.403097349  0.060709082
## factor(street)74     -0.218550952  0.088371435
## factor(street)75     -0.248696424  0.136046051
## factor(street)77     -0.280334541  0.213787226
## factor(street)78     -0.220793273  0.117282571
## factor(street)79     -0.015970142  0.629391879
## factor(street)80     -0.123629404  0.146876525
## factor(street)81     -0.353958842  0.291850628
## factor(street)82     -0.338480016  0.404040619
## factor(street)83     -0.366653294  0.128957376
## factor(street)84     -0.264697068  0.228955552
## factor(street)85     -0.103975707  0.234081672
## factor(street)86     -0.223830031  0.089751576
## factor(street)88     -0.306489334  0.111358402
## factor(street)89     -0.235677897  0.228499698
## factor(street)91     -0.581978352  0.160316839
## factor(street)93     -0.317038251  0.176746280
## factor(street)94     -0.256662975  0.182594029
## factor(street)96     -0.285865344  0.207981007
## factor(street)98     -0.140120013  0.244869511
```

```
## factor(street)99      0.011638952  0.451235725
## factor(street)100    -0.134841581  0.283514602
## factor(street)101    -0.022103126  0.416693380
## factor(street)102    -0.458135869  0.287420697
## factor(street)103    -0.133357329  0.305697285
## factor(street)105    -0.173744662  0.405789681
## factor(street)106     0.208202515  0.592710454
## factor(street)107    -0.258215971  0.386914375
## factor(street)109    -0.246305937  0.217237749
## factor(street)110    -0.148956449  0.430484181
## factor(street)111    -0.288370726  0.291119746
## factor(street)112    -0.047345774  0.416920807
## factor(street)113    -0.165515682  0.329073053
## factor(street)115     0.100168978  0.484600572
## factor(street)117     0.183792566  0.930861134
## factor(street)118    -0.338111585  0.403854476
## factor(street)119    -0.242752815  0.220826216
## factor(street)120    -0.277948086  0.160809473
## factor(street)121    -0.053838562  0.410150244
## factor(street)122    -0.483297353  0.161879117
## factor(street)124    -0.160507269  0.239476519
## factor(street)125     0.129772158  0.469420361
## factor(street)126    -0.369669353  0.162632067
## factor(street)127    -0.192606575  0.339611550
## factor(street)128    -0.046920744  0.337579137
## factor(street)129     0.005766076  0.748372233
## factor(street)130    -0.055791071  0.407926105
## factor(street)131     0.043088985  0.508187625
## factor(street)132     0.060463491  0.408833077
## factor(street)133    -0.006814659  0.638546396
## factor(street)134    -0.257453192  0.236458202
## factor(street)136    -0.051242828  0.412924533
## factor(street)137    -0.115889395  0.626076665
## factor(street)138    -0.079848140  0.322082135
## factor(street)147    -0.178515039  0.563951162
## factor(street)148    -0.216549273  0.428581073
## factor(street)149     0.061812344  0.501594117
## factor(street)151    -0.634710424  0.107888733
## factor(street)152    -0.246371026  0.495602836
## factor(street)153    -0.398755856  0.343294228
## factor(street)154    -0.268342359  0.149611460
## factor(street)155    -0.615889391  0.126076670
## factor(street)156    -0.175256243  0.566741356
## factor(street)157    -0.174433509  0.472048576
## factor(street)158     0.108086031  0.687980363
## factor(street)160    -0.321518416  0.078536853
## factor(street)163    -0.325116847  0.169099059
## factor(street)164    -0.387605634  0.191856077
## factor(street)165    -0.313546865  0.332758618
## factor(street)166    -0.518667167  0.223298894
## factor(street)168    -0.220932699  0.522853239
## factor(street)170    -0.181265299  0.166402824
## factor(street)171    -0.312603812  0.218408889
## factor(street)172    -0.257109942  0.206456687
```

```
## factor(street)175    0.065773523  0.436649079
## factor(street)179   -0.319073538  0.326451869
## factor(street)180   -0.247751445  0.052576852
## factor(street)182   -0.411090963  0.119998345
## factor(street)183   -0.236450810  0.070427870
## factor(street)185   -0.327575760  0.203489758
## factor(street)186   -0.228308738  0.235432101
## factor(street)187   -0.189299289  0.169603491
## factor(street)188   -0.247936442  0.190778575
## factor(street)189   -0.205313494  0.095091357
## factor(street)190   -0.221194003  0.243575701
## factor(street)191   -0.234866809  0.203851216
## factor(street)192   -0.242197792  0.032508536
## factor(street)193   -0.169836158  0.293742873
## factor(street)196   -0.492991107  0.038116157
## factor(street)197   -0.299833683  0.442181002
## factor(street)198   -0.090343656  0.257332750
## factor(street)200   -0.291290470  0.202745025
## factor(street)202   -0.283121236  0.296590303
## factor(street)203   -0.472794373  0.020755674
## factor(street)206   -0.446731321 -0.028806279
## factor(street)207   -0.559470042  0.085878531
## factor(street)208   -0.325897991  0.092027051
## factor(street)209   -0.494943945 -0.124294050
## factor(street)210   -0.438796824 -0.020511460
## factor(street)213   -0.150119124  0.234876398
## factor(street)215   -0.186799241  0.068938972
## factor(street)216   -0.748650049 -0.164182437
## factor(street)217   -0.411310484  0.052370658
## factor(street)220   -0.452070891  0.041514117
## factor(street)221   -0.369993982  0.068730897
## factor(street)222   -0.464153357  0.115459865
## factor(street)223   -0.411327530  0.052282580
## factor(street)225   -0.310047017  0.129546835
## factor(street)227   -0.336336923  0.243123305
## factor(street)228   -0.321389546  0.420776900
## factor(street)229   -0.230168198  0.303361954
## factor(street)230   -0.236791862  0.256909314
## factor(street)232   -0.245215105  0.400813450
## factor(street)233   -0.139835168  0.244700317
## factor(street)235   -0.401214039  0.179171216
## factor(street)236   -0.173935827  0.147040985
## factor(street)238   -0.327744353  0.111422054
## factor(street)240   -0.586779634 -0.007359215
## factor(street)241   -0.426366742  0.219071817
## factor(street)242   -0.331378416  0.410579101
## factor(street)243   -0.164754711  0.481326373
## factor(street)244   -0.088596108  0.270131330
## factor(street)246   -0.403128449  0.128626822
## factor(street)247   -0.188122102  0.182516938
## factor(street)248   -0.145197253  0.133872786
## factor(street)249   -0.082139307  0.255692768
## factor(street)250   -0.076615517  0.167452600
## factor(street)253   -0.055494270  0.273549593
```

```
## factor(street)254   -0.145953396  0.348186951
## factor(street)255   -0.146488976  0.201684259
## factor(street)256   -0.101704742  0.151301086
## factor(street)257   -0.264125198  0.056893963
## factor(street)258   -0.053751164  0.247123957
## factor(street)259    0.330308809  1.072811750
## factor(street)260   -0.156899801  0.092595011
## factor(street)261   -0.275647921  0.188561284
## factor(street)262   -0.183802251  0.255091764
## factor(street)263   -0.137596776  0.191375192
```

Narrative: Adding street fixed effects accounts for unobserved heterogeneity across streets. The confidence interval changes slightly, reflecting the improved precision.

3.4 Test for block fixed effects

Recall that the authors described their experiment as "stratified at the street level," which is a synonym for blocking by street. Does including these block fixed effects change the standard errors of the estimates *very much*? Conduct the appropriate test for the inclusion of these block fixed effects, and interpret them in the context of the other variables in the regression.

```r
mod_pretreat <- lm(avg_bins_treat ~ bin + base_avg_bins_treat, data = d_clean)
mod_fixed_effects <- lm(avg_bins_treat ~ bin + base_avg_bins_treat + factor(street), data = d_clean)

# Now perform the ANOVA test
test_fixed_effects <- anova(mod_pretreat, mod_fixed_effects)
test_fixed_effects
```

```
## Analysis of Variance Table
##
## Model 1: avg_bins_treat ~ bin + base_avg_bins_treat
## Model 2: avg_bins_treat ~ bin + base_avg_bins_treat + factor(street)
##   Res.Df    RSS  Df Sum of Sq      F    Pr(>F)
## 1   1779 196.78
## 2   1600 167.50 179    29.278 1.5624 9.514e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer: The inclusion of block fixed effects significantly improves model fit, as indicated by the ANOVA test.

3.5 Feature (no) cell phone

Perhaps having a cell phone helps explain the level of recycling behavior. Instead of "has cell phone," we find it easier to interpret the coefficient if we define the variable " no cell phone." Give the R command to define this new variable, which equals one minus the "has cell phone" variable in the authors' data set. Use "no cell phone" instead of "has cell phone" in subsequent regressions with this dataset.

```r
## rather than letting this feature engineering persist here -- which is
## bad practice because it requires that you keep in mind what code you
## have, and what code you have not run
##
## instead, move this recode up to the top of your data loading in this file
## so that it runs everytime that you load your data
```

Now add "no cell phone" as a covariate to the previous regression. Provide a 95% confidence interval for the treatment effect. Explain why this confidence interval does not differ much from the previous one.

```r
d_clean$no_cellphone <- 1 - d_clean$havecell

mod_cellphone <- lm(avg_bins_treat ~ bin + base_avg_bins_treat + no_cellphone, data = d_clean)
coeftest(mod_cellphone, vcov = vcovHC(mod_cellphone, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                      Estimate Std. Error t value  Pr(>|t|)
## (Intercept)          0.372153   0.021846 17.0354 < 2.2e-16 ***
## bin                  0.125163   0.017104  7.3177 3.801e-13 ***
## base_avg_bins_treat  0.389530   0.029632 13.1455 < 2.2e-16 ***
## no_cellphone        -0.049013   0.015625 -3.1368  0.001736 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
confint(mod_cellphone)
```

```
##                           2.5 %      97.5 %
## (Intercept)          0.34231221  0.40199333
## bin                  0.09251751  0.15780820
## base_avg_bins_treat  0.36323580  0.41582481
## no_cellphone        -0.08038993 -0.01763689
```

Answer: The confidence interval remains mostly unchanged since cell phone ownership is not strongly correlated with treatment.

3.6 Add the sms treatment

Now let's add in the SMS treatment. Re-run the previous regression with "any SMS" included. You should get the same results as in Table 4A. Provide a 95% confidence interval for the treatment effect of the recycling bin. Explain why this confidence interval does not differ much from the previous one.

```r
mod_sms <- lm(avg_bins_treat ~ bin + base_avg_bins_treat + no_cellphone + sms, data = d_clean)
coeftest(mod_sms, vcov = vcovHC(mod_sms, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                      Estimate Std. Error t value  Pr(>|t|)
## (Intercept)          0.389190   0.024392 15.9558 < 2.2e-16 ***
## bin                  0.125185   0.017098  7.3216 3.698e-13 ***
## base_avg_bins_treat  0.388267   0.029578 13.1269 < 2.2e-16 ***
## no_cellphone        -0.065139   0.019743 -3.2993 0.0009884 ***
## sms                 -0.030867   0.021243 -1.4530 0.1463923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
confint(mod_sms)
```

```
##                           2.5 %       97.5 %
## (Intercept)          0.35199872  0.426381152
## bin                  0.09255130  0.157818847
## base_avg_bins_treat  0.36193004  0.414603519
## no_cellphone        -0.10289827 -0.027379704
## sms                 -0.07110934  0.009374968
```

Answer: The confidence interval for the recycling bin treatment remains stable, suggesting SMS treatment does not strongly interact with bin provision.

3.7 Reproduce Table 4B, Column (2)

Now reproduce the results of column 2 in Table 4B, estimating separate treatment effects for the two types of SMS treatments and the two types of recycling-bin treatments. Provide a 95% confidence interval for the effect of the unadorned recycling bin. Explain how your answer differs from that in question 3.6, and explain why you think it differs.

```
mod_full <- lm(avg_bins_treat ~ bin_s + bin_g + sms_p + sms_g + base_avg_bins_treat + no_cellphone + fa
coeftest(mod_full, vcov = vcovHC(mod_full, type = "HC1"))
```

```
##
## t test of coefficients:
##
##                         Estimate  Std. Error t value  Pr(>|t|)
## (Intercept)           3.8494e-01  3.7717e-02 10.2060 < 2.2e-16 ***
## bin_s                 1.2781e-01  2.2100e-02  5.7833 8.796e-09 ***
## bin_g                 1.0319e-01  2.3216e-02  4.4449 9.401e-06 ***
## sms_p                -8.0412e-03  2.6185e-02 -0.3071 0.7588184
## sms_g                 1.9707e-02  2.5774e-02  0.7646 0.4446187
## base_avg_bins_treat   3.7385e-01  2.6974e-02 13.8595 < 2.2e-16 ***
## no_cellphone         -4.6383e-02  2.1440e-02 -2.1635 0.0306543 *
## factor(street)2      -1.0068e-01  5.7106e-02 -1.7630 0.0780906 .
## factor(street)3      -5.4913e-02  9.9369e-02 -0.5526 0.5806022
## factor(street)4      -1.3037e-01  6.6878e-02 -1.9494 0.0514203 .
## factor(street)5      -1.0415e-01  9.8476e-02 -1.0576 0.2904096
## factor(street)6      -5.7119e-02  1.0981e-01 -0.5202 0.6030111
## factor(street)7      -7.5412e-02  1.2748e-01 -0.5916 0.5542245
## factor(street)8      -5.7272e-02  5.6863e-02 -1.0072 0.3139945
## factor(street)9      -9.2053e-02  9.8040e-02 -0.9389 0.3479096
## factor(street)10     -3.1455e-02  8.0283e-02 -0.3918 0.6952567
## factor(street)11     -8.2240e-02  1.0792e-01 -0.7621 0.4461419
## factor(street)15     -3.4969e-02  9.9538e-02 -0.3513 0.7253988
## factor(street)17      3.1179e-02  9.3284e-02  0.3342 0.7382464
## factor(street)20     -2.3101e-01  6.9540e-02 -3.3219 0.0009142 ***
## factor(street)21     -6.4841e-02  7.8754e-02 -0.8233 0.4104386
## factor(street)22     -1.1723e-01  6.7796e-02 -1.7292 0.0839724 .
## factor(street)23     -1.5904e-01  5.6113e-02 -2.8343 0.0046506 **
## factor(street)26     -7.9013e-03  7.5359e-02 -0.1048 0.9165097
## factor(street)32      1.7029e-01  2.0137e-01  0.8457 0.3978603
## factor(street)37      1.4667e-01  1.7136e-01  0.8559 0.3921795
## factor(street)38     -8.6016e-02  8.3199e-02 -1.0339 0.3013601
## factor(street)40     -5.4383e-02  2.4518e-01 -0.2218 0.8244920
## factor(street)41      1.3738e-02  1.5542e-01  0.0884 0.9295728
## factor(street)42     -1.1676e-01  1.1256e-01 -1.0373 0.2997581
## factor(street)43      3.7217e-02  1.3279e-01  0.2803 0.7793108
## factor(street)44      1.1159e-01  1.1148e-01  1.0010 0.3169843
## factor(street)45     -4.4788e-02  9.8287e-02 -0.4557 0.6486742
## factor(street)46      1.7746e-03  9.3375e-02  0.0190 0.9848393
## factor(street)47      4.0174e-01  2.4702e-01  1.6263 0.1040759
## factor(street)53      1.0397e-02  1.5304e-01  0.0679 0.9458462
## factor(street)58     -3.2325e-02  6.0479e-02 -0.5345 0.5930863
## factor(street)60      2.6221e-02  7.0758e-02  0.3706 0.7110035
```

```
## factor(street)61     -5.2121e-03  1.2446e-01 -0.0419 0.9666024
## factor(street)62      4.0303e-02  1.2160e-01  0.3314 0.7403676
## factor(street)63     -3.1652e-02  8.7545e-02 -0.3615 0.7177390
## factor(street)64      3.8909e-02  9.2468e-02  0.4208 0.6739737
## factor(street)66     -5.1151e-02  5.3454e-02 -0.9569 0.3387588
## factor(street)67     -8.9142e-03  1.4625e-01 -0.0610 0.9514064
## factor(street)68      8.2547e-02  1.0949e-01  0.7539 0.4510127
## factor(street)69      6.6945e-02  7.9799e-02  0.8389 0.4016426
## factor(street)70     -2.2194e-02  8.2918e-02 -0.2677 0.7889916
## factor(street)72     -4.5950e-02  8.1122e-02 -0.5664 0.5711830
## factor(street)73     -1.8062e-01  5.4939e-02 -3.2877 0.0010322 **
## factor(street)74     -7.3581e-02  7.0023e-02 -1.0508 0.2935071
## factor(street)75     -5.0793e-02  5.0009e-02 -1.0157 0.3099334
## factor(street)77     -3.8744e-02  1.5572e-01 -0.2488 0.8035394
## factor(street)78     -6.3666e-02  8.0747e-02 -0.7885 0.4305466
## factor(street)79      3.1723e-01  4.1750e-01  0.7598 0.4474691
## factor(street)80      1.0752e-02  5.1905e-02  0.2071 0.8359248
## factor(street)81     -2.8831e-02  1.1314e-01 -0.2548 0.7988879
## factor(street)82      3.8088e-02  1.6328e-01  0.2333 0.8155774
## factor(street)83     -1.1813e-01  1.1954e-01 -0.9882 0.3232061
## factor(street)84     -1.7787e-03  6.5843e-02 -0.0270 0.9784522
## factor(street)85      6.3635e-02  7.6080e-02  0.8364 0.4030380
## factor(street)86     -7.6893e-02  7.3244e-02 -1.0498 0.2939644
## factor(street)88     -9.9295e-02  1.0086e-01 -0.9845 0.3250062
## factor(street)89     -9.8977e-03  9.4219e-02 -0.1050 0.9163498
## factor(street)91     -2.3929e-01  1.2475e-01 -1.9181 0.0552796 .
## factor(street)93     -7.1193e-02  8.6510e-02 -0.8229 0.4106628
## factor(street)94     -5.1011e-02  7.9874e-02 -0.6386 0.5231469
## factor(street)96     -3.6466e-02  1.5782e-01 -0.2311 0.8172942
## factor(street)98      3.9868e-02  8.3190e-02  0.4792 0.6318383
## factor(street)99      2.3238e-01  9.0081e-02  2.5797 0.0099787 **
## factor(street)100     6.9302e-02  1.3948e-01  0.4969 0.6193448
## factor(street)101     2.0898e-01  1.1255e-01  1.8568 0.0635232 .
## factor(street)102    -5.2905e-02  1.4371e-01 -0.3681 0.7128289
## factor(street)103     8.7214e-02  1.3203e-01  0.6605 0.5089998
## factor(street)105     1.1426e-01  8.3272e-02  1.3722 0.1702030
## factor(street)106     3.8918e-01  1.2962e-01  3.0026 0.0027189 **
## factor(street)107     4.5574e-02  1.5735e-01  0.2896 0.7721332
## factor(street)109    -2.4444e-02  1.1901e-01 -0.2054 0.8372900
## factor(street)110     1.2935e-01  1.9041e-01  0.6793 0.4970346
## factor(street)111    -2.3856e-02  1.5015e-01 -0.1589 0.8737833
## factor(street)112     1.8502e-01  8.3203e-02  2.2237 0.0263104 *
## factor(street)113     5.6358e-02  1.6157e-01  0.3488 0.7272745
## factor(street)115     2.7738e-01  1.3385e-01  2.0723 0.0383991 *
## factor(street)117     5.6607e-01  4.1141e-01  1.3759 0.1690384
## factor(street)118     1.2547e-02  2.6536e-01  0.0473 0.9622920
## factor(street)119    -2.0902e-02  1.0640e-01 -0.1965 0.8442794
## factor(street)120    -8.0370e-02  1.0071e-01 -0.7980 0.4249648
## factor(street)121     1.6991e-01  1.6828e-01  1.0097 0.3128009
## factor(street)122    -1.6869e-01  1.1099e-01 -1.5198 0.1287510
## factor(street)124     2.7000e-02  1.0843e-01  0.2490 0.8033844
## factor(street)125     2.9550e-01  1.4627e-01  2.0202 0.0435261 *
## factor(street)126    -1.1833e-01  1.2703e-01 -0.9315 0.3517232
## factor(street)127     6.7742e-02  3.3354e-01  0.2031 0.8390828
```

```
## factor(street)128    1.4648e-01  1.0804e-01  1.3558 0.1753603
## factor(street)129    3.7679e-01  4.5318e-01  0.8314 0.4058509
## factor(street)130    1.7459e-01  1.3634e-01  1.2805 0.2005460
## factor(street)131    2.7003e-01  1.4931e-01  1.8086 0.0707018 .
## factor(street)132    2.3389e-01  1.0112e-01  2.3131 0.0208435 *
## factor(street)133    3.1479e-01  1.1806e-01  2.6663 0.0077466 **
## factor(street)134   -8.5634e-03  1.1979e-01 -0.0715 0.9430190
## factor(street)136    1.6015e-01  1.3733e-01  1.1662 0.2437240
## factor(street)137    2.4202e-01  1.0442e-01  2.3177 0.0205919 *
## factor(street)138    1.0922e-01  1.6452e-01  0.6639 0.5068790
## factor(street)147    1.8979e-01  5.9383e-02  3.1961 0.0014204 **
## factor(street)148    1.0298e-01  2.1537e-01  0.4782 0.6326049
## factor(street)149    2.6780e-01  1.1486e-01  2.3314 0.0198541 *
## factor(street)151   -2.7655e-01  1.8189e-01 -1.5204 0.1286122
## factor(street)152    1.4181e-01  2.2670e-01  0.6256 0.5316874
## factor(street)153   -1.6109e-02  1.0320e-01 -0.1561 0.8759785
## factor(street)154   -5.7247e-02  1.1474e-01 -0.4989 0.6178982
## factor(street)155   -2.7076e-01  1.7344e-01 -1.5611 0.1186943
## factor(street)156    2.1563e-01  1.7429e-01  1.2372 0.2162111
## factor(street)157    1.3938e-01  7.0657e-02  1.9727 0.0487065 *
## factor(street)158    3.8918e-01  1.9865e-01  1.9592 0.0502681 .
## factor(street)160   -1.1861e-01  8.2203e-02 -1.4428 0.1492604
## factor(street)163   -1.0772e-01  8.1235e-02 -1.3260 0.1850158
## factor(street)164   -9.0318e-02  5.9293e-02 -1.5232 0.1278966
## factor(street)165    1.3261e-02  1.1031e-01  0.1202 0.9043312
## factor(street)166   -1.3709e-01  1.2384e-01 -1.1069 0.2684935
## factor(street)168    1.6710e-01  3.6593e-01  0.4566 0.6479879
## factor(street)170   -5.8723e-03  6.4558e-02 -0.0910 0.9275337
## factor(street)171   -3.6785e-02  9.9005e-02 -0.3716 0.7102773
## factor(street)172   -6.3041e-03  4.7542e-02 -0.1326 0.8945256
## factor(street)175    2.6195e-01  1.6428e-01  1.5945 0.1110145
## factor(street)179   -2.2306e-02  1.4010e-01 -0.1592 0.8735201
## factor(street)180   -9.2607e-02  6.7208e-02 -1.3779 0.1684256
## factor(street)182   -1.4619e-01  1.1130e-01 -1.3135 0.1892079
## factor(street)183   -7.2429e-02  7.6485e-02 -0.9470 0.3437972
## factor(street)185   -6.2725e-02  8.9389e-02 -0.7017 0.4829674
## factor(street)186   -4.7802e-03  8.7218e-02 -0.0548 0.9562987
## factor(street)187    6.1195e-06  7.3092e-02  0.0001 0.9999332
## factor(street)188   -2.9999e-02  4.9041e-02 -0.6117 0.5408140
## factor(street)189   -5.3966e-02  6.3721e-02 -0.8469 0.3971659
## factor(street)190    2.0221e-02  1.1204e-01  0.1805 0.8567972
## factor(street)191   -2.1182e-02  9.4782e-02 -0.2235 0.8231931
## factor(street)192   -1.0166e-01  6.5589e-02 -1.5500 0.1213357
## factor(street)193    7.8284e-02  1.0938e-01  0.7157 0.4742622
## factor(street)196   -2.3952e-01  1.0869e-01 -2.2038 0.0276834 *
## factor(street)197    4.5310e-02  3.3016e-01  0.1372 0.8908613
## factor(street)198    8.9852e-02  1.0940e-01  0.8213 0.4115752
## factor(street)200   -3.4842e-02  6.0629e-02 -0.5747 0.5655853
## factor(street)202    2.4680e-02  1.3434e-01  0.1837 0.8542645
## factor(street)203   -2.4234e-01  8.1305e-02 -2.9806 0.0029202 **
## factor(street)206   -2.3951e-01  9.4022e-02 -2.5474 0.0109457 *
## factor(street)207   -2.3234e-01  5.1398e-02 -4.5204 6.629e-06 ***
## factor(street)208   -1.0606e-01  1.0602e-01 -1.0004 0.3172925
## factor(street)209   -3.1585e-01  8.6698e-02 -3.6431 0.0002780 ***
```

```
## factor(street)210   -2.0589e-01  8.3721e-02 -2.4593 0.0140270 *
## factor(street)213    5.1422e-02  8.2544e-02  0.6230 0.5333924
## factor(street)215   -5.3807e-02  7.2682e-02 -0.7403 0.4592239
## factor(street)216   -4.4948e-01  3.4897e-01 -1.2880 0.1979204
## factor(street)217   -1.8981e-01  1.1190e-01 -1.6962 0.0900375 .
## factor(street)220   -1.9819e-01  1.3550e-01 -1.4626 0.1437661
## factor(street)221   -1.6433e-01  9.2856e-02 -1.7697 0.0769693 .
## factor(street)222   -1.5236e-01  1.2673e-01 -1.2023 0.2294407
## factor(street)223   -1.7726e-01  1.2074e-01 -1.4681 0.1422760
## factor(street)225   -9.0615e-02  1.1122e-01 -0.8147 0.4153623
## factor(street)227   -6.2553e-02  1.5160e-01 -0.4126 0.6799405
## factor(street)228    4.7572e-02  1.8568e-01  0.2562 0.7978318
## factor(street)229    5.0362e-02  1.4375e-01  0.3503 0.7261275
## factor(street)230    1.2024e-02  8.3786e-02  0.1435 0.8859076
## factor(street)232    8.2628e-02  2.3091e-01  0.3578 0.7205070
## factor(street)233    4.8806e-02  1.3316e-01  0.3665 0.7140243
## factor(street)235   -1.2687e-01  1.5003e-01 -0.8456 0.3978975
## factor(street)236   -5.4007e-04  7.1785e-02 -0.0075 0.9939982
## factor(street)238   -9.8937e-02  9.0470e-02 -1.0936 0.2743043
## factor(street)240   -3.0366e-01  6.6887e-02 -4.5399 6.050e-06 ***
## factor(street)241   -1.1921e-01  1.9142e-01 -0.6228 0.5335154
## factor(street)242    5.0223e-02  6.0467e-02  0.8306 0.4063347
## factor(street)243    1.3872e-01  2.6411e-01  0.5252 0.5994850
## factor(street)244    1.0885e-01  1.1227e-01  0.9695 0.3324252
## factor(street)246   -1.3209e-01  1.0379e-01 -1.2726 0.2033322
## factor(street)247   -2.0474e-03  5.3529e-02 -0.0382 0.9694947
## factor(street)248    3.5715e-04  7.5041e-02  0.0048 0.9962031
## factor(street)249    9.0261e-02  1.2106e-01  0.7456 0.4560081
## factor(street)250    4.9569e-02  5.0438e-02  0.9828 0.3258622
## factor(street)253    1.2179e-01  8.5563e-02  1.4234 0.1548157
## factor(street)254    1.1358e-01  1.1441e-01  0.9928 0.3209809
## factor(street)255    3.7027e-02  1.0236e-01  0.3617 0.7175933
## factor(street)256    3.4617e-02  6.3663e-02  0.5438 0.5866821
## factor(street)257   -1.0495e-01  7.1714e-02 -1.4635 0.1435259
## factor(street)258    9.3953e-02  5.5043e-02  1.7069 0.0880348 .
## factor(street)259    7.2954e-01  2.4355e-01  2.9954 0.0027830 **
## factor(street)260   -3.4424e-02  7.5290e-02 -0.4572 0.6475721
## factor(street)261   -3.5855e-02  7.7170e-02 -0.4646 0.6422673
## factor(street)262    4.9097e-02  1.1049e-01  0.4444 0.6568393
## factor(street)263    2.7669e-02  9.5631e-02  0.2893 0.7723666
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
confint(mod_full)
```

```
##                            2.5 %        97.5 %
## (Intercept)          0.317676405  0.4522102273
## bin_s                0.084197488  0.1714282957
## bin_g                0.060256268  0.1461241648
## sms_p               -0.057149905  0.0410676007
## sms_g               -0.029716858  0.0691310914
## base_avg_bins_treat  0.345776609  0.4019277468
## no_cellphone        -0.086354787 -0.0064121324
## factor(street)2     -0.300401958  0.0990444661
## factor(street)3     -0.263668881  0.1538424628
```

```
## factor(street)4      -0.349690777  0.0889426677
## factor(street)5      -0.323220470  0.1149294184
## factor(street)6      -0.288762122  0.1745245026
## factor(street)7      -0.365252512  0.2144275994
## factor(street)8      -0.236303212  0.1217587901
## factor(street)9      -0.283977648  0.0998725415
## factor(street)10     -0.188001694  0.1250916934
## factor(street)11     -0.274562350  0.1100832733
## factor(street)15     -0.357770677  0.2878323219
## factor(street)17     -0.234131257  0.2964890072
## factor(street)20     -0.462630077  0.0006173745
## factor(street)21     -0.331021721  0.2013394775
## factor(street)22     -0.252316559  0.0178562040
## factor(street)23     -0.300918668 -0.0171642096
## factor(street)26     -0.378619688  0.3628171807
## factor(street)32     -0.200931455  0.5415158464
## factor(street)37     -0.143766778  0.4370988114
## factor(street)38     -0.375615390  0.2035832484
## factor(street)40     -0.425459147  0.3166939599
## factor(street)41     -0.275920903  0.3033977772
## factor(street)42     -0.488220648  0.2547036989
## factor(street)43     -0.252404643  0.3268394786
## factor(street)44     -0.120563432  0.3437473845
## factor(street)45     -0.367006723  0.2774297462
## factor(street)46     -0.263451001  0.2670002355
## factor(street)47      0.078671434  0.7248043357
## factor(street)53     -0.360395005  0.3811886796
## factor(street)58     -0.217467274  0.1528182238
## factor(street)60     -0.142513299  0.1949555860
## factor(street)61     -0.376589976  0.3661657559
## factor(street)62     -0.206777488  0.2873826782
## factor(street)63     -0.263244522  0.1999410764
## factor(street)64     -0.226312983  0.3041300532
## factor(street)66     -0.297609626  0.1953077075
## factor(street)67     -0.228095062  0.2102666769
## factor(street)68     -0.148946281  0.3140402778
## factor(street)69     -0.133083722  0.2669734567
## factor(street)70     -0.201490325  0.1571018977
## factor(street)72     -0.245666746  0.1537675449
## factor(street)73     -0.412228579  0.0509857356
## factor(street)74     -0.226877606  0.0797160056
## factor(street)75     -0.243089843  0.1415030348
## factor(street)77     -0.285458313  0.2079699237
## factor(street)78     -0.232659160  0.1053275138
## factor(street)79     -0.005527356  0.6399905306
## factor(street)80     -0.124307265  0.1458107705
## factor(street)81     -0.351331078  0.2936690865
## factor(street)82     -0.333222889  0.4093996289
## factor(street)83     -0.365588801  0.1293282425
## factor(street)84     -0.248505870  0.2449485574
## factor(street)85     -0.105345709  0.2326165086
## factor(street)86     -0.233569491  0.0797838552
## factor(street)88     -0.308047100  0.1094565756
## factor(street)89     -0.241723296  0.2219279895
```

```
## factor(street)91     -0.610429347  0.1318501405
## factor(street)93     -0.317735687  0.1753501219
## factor(street)94     -0.270634343  0.1686120949
## factor(street)96     -0.282981708  0.2100492148
## factor(street)98     -0.152880240  0.2326152503
## factor(street)99      0.012913310  0.4518410301
## factor(street)100    -0.139548570  0.2781528842
## factor(street)101    -0.010391973  0.4283573451
## factor(street)102    -0.425644157  0.3198345333
## factor(street)103    -0.132021321  0.3064493173
## factor(street)105    -0.175036678  0.4035626316
## factor(street)106     0.196904073  0.5814496483
## factor(street)107    -0.277122870  0.3682711529
## factor(street)109    -0.256240889  0.2073530274
## factor(street)110    -0.160729995  0.4194219806
## factor(street)111    -0.313769697  0.2660583559
## factor(street)112    -0.046953260  0.4169843233
## factor(street)113    -0.191511464  0.3042284154
## factor(street)115     0.085160265  0.4695939947
## factor(street)117     0.192831512  0.9393176012
## factor(street)118    -0.358327528  0.3834224170
## factor(street)119    -0.252977986  0.2111739136
## factor(street)120    -0.312413579  0.1516733760
## factor(street)121    -0.062198185  0.4020230304
## factor(street)122    -0.490944579  0.1535597103
## factor(street)124    -0.173361894  0.2273628682
## factor(street)125     0.125879835  0.4651215714
## factor(street)126    -0.384830463  0.1481761749
## factor(street)127    -0.198083403  0.3335666171
## factor(street)128    -0.045722426  0.3386807008
## factor(street)129     0.006000045  0.7475788647
## factor(street)130    -0.057065965  0.4062527598
## factor(street)131     0.037529469  0.5025404182
## factor(street)132     0.059719924  0.4080632842
## factor(street)133    -0.007926898  0.6374992529
## factor(street)134    -0.255106491  0.2379796100
## factor(street)136    -0.071928511  0.3922258683
## factor(street)137    -0.128627966  0.6126744224
## factor(street)138    -0.092138774  0.3105697737
## factor(street)147    -0.181120453  0.5607091214
## factor(street)148    -0.219545823  0.4255099932
## factor(street)149     0.047145729  0.4884551907
## factor(street)151    -0.647523968  0.0944281569
## factor(street)152    -0.228791003  0.5124180943
## factor(street)153    -0.386942371  0.3547247261
## factor(street)154    -0.265926328  0.1514320793
## factor(street)155    -0.641560564  0.1000454928
## factor(street)156    -0.155064795  0.5863180770
## factor(street)157    -0.184197094  0.4629634144
## factor(street)158     0.099622941  0.6787441562
## factor(street)160    -0.318324946  0.0811128788
## factor(street)163    -0.355504229  0.1400631313
## factor(street)164    -0.379730491  0.1990952368
## factor(street)165    -0.309545604  0.3360668139
```

```
## factor(street)166   -0.507777939  0.2336062769
## factor(street)168   -0.204397388  0.5385965361
## factor(street)170   -0.179500790  0.1677561300
## factor(street)171   -0.302253490  0.2286827593
## factor(street)172   -0.238151551  0.2255433099
## factor(street)175    0.076672606  0.4472360001
## factor(street)179   -0.345004747  0.3003923207
## factor(street)180   -0.242567901  0.0573543107
## factor(street)182   -0.411393884  0.1190186273
## factor(street)183   -0.225850672  0.0809925214
## factor(street)185   -0.328073651  0.2026240659
## factor(street)186   -0.236402738  0.2268423584
## factor(street)187   -0.179338746  0.1793509846
## factor(street)188   -0.249017358  0.1890188909
## factor(street)189   -0.204048928  0.0961160436
## factor(street)190   -0.211890265  0.2523322633
## factor(street)191   -0.240213504  0.1978503717
## factor(street)192   -0.238801694  0.0354742043
## factor(street)193   -0.153480591  0.3100489609
## factor(street)196   -0.504787883  0.0257437104
## factor(street)197   -0.325518611  0.4161393407
## factor(street)198   -0.083749104  0.2634527509
## factor(street)200   -0.281560513  0.2118756046
## factor(street)202   -0.265095101  0.3144549542
## factor(street)203   -0.489145142  0.0044668004
## factor(street)206   -0.448153308 -0.0308706373
## factor(street)207   -0.554667165  0.0899907952
## factor(street)208   -0.315042138  0.1029235857
## factor(street)209   -0.500932470 -0.1307677725
## factor(street)210   -0.415200560  0.0034134526
## factor(street)213   -0.141026275  0.2438710308
## factor(street)215   -0.181541173  0.0739275784
## factor(street)216   -0.741403624 -0.1575577624
## factor(street)217   -0.421398422  0.0417827090
## factor(street)220   -0.444644722  0.0482613826
## factor(street)221   -0.383556167  0.0549025041
## factor(street)222   -0.442036646  0.1373131116
## factor(street)223   -0.408758248  0.0542428103
## factor(street)225   -0.310129476  0.1289003623
## factor(street)227   -0.352170466  0.2270644133
## factor(street)228   -0.323148626  0.4182922399
## factor(street)229   -0.216128755  0.3168535486
## factor(street)230   -0.234415286  0.2584631715
## factor(street)232   -0.240008831  0.4052653444
## factor(street)233   -0.143180786  0.2407931864
## factor(street)235   -0.416938465  0.1631995307
## factor(street)236   -0.161010953  0.1599308065
## factor(street)238   -0.318241262  0.1203681572
## factor(street)240   -0.593096077 -0.0142265190
## factor(street)241   -0.441839363  0.2034148785
## factor(street)242   -0.320462320  0.4209080898
## factor(street)243   -0.184483052  0.4619261762
## factor(street)244   -0.070605090  0.2883036157
## factor(street)246   -0.397937113  0.1337579773
```

```
## factor(street)247   -0.187083748  0.1829890022
## factor(street)248   -0.138998822  0.1397131282
## factor(street)249   -0.078649481  0.2591722479
## factor(street)250   -0.072468962  0.1716074506
## factor(street)253   -0.042657688  0.2862389223
## factor(street)254   -0.133555716  0.3607243116
## factor(street)255   -0.136899067  0.2109538521
## factor(street)256   -0.091839627  0.1610743313
## factor(street)257   -0.265281934  0.0553743768
## factor(street)258   -0.056311345  0.2442177609
## factor(street)259    0.358403369  1.1006779924
## factor(street)260   -0.158991483  0.0901428119
## factor(street)261   -0.267748180  0.1960386133
## factor(street)262   -0.170227595  0.2684221804
## factor(street)263   -0.136606258  0.1919441794
```

Answer: The separate treatment effects for different bin and SMS conditions align with Table 4B, Column 2. Differences arise due to the inclusion of fixed effects and additional covariates.

4 A Final Practice Problem

4.1 Simple treatment effect of Zmapp

Without using any covariates, answer this question with regression: What is the estimated effect of ZMapp (with standard error in parentheses) on whether someone was dehydrated on day 14? What is the p-value associated with this estimate?

```r
zmapp_1 <- glm(dehydrated_day14 ~ treat_zmapp, data = d, family = binomial)
summary(zmapp_1)
```

```
##
## Call:
## glm(formula = dehydrated_day14 ~ treat_zmapp, family = binomial,
##     data = d)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.7148     0.3621   4.736 2.18e-06 ***
## treat_zmapp  -1.2685     0.4833  -2.625  0.00868 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 112.47  on 99  degrees of freedom
## Residual deviance: 105.24  on 98  degrees of freedom
## AIC: 109.24
##
## Number of Fisher Scoring iterations: 4
```

Answer: The estimated effect of ZMapp on dehydration on day 14 is -1.2685 with a standard error of (0.4833). The p-value for this estimate is 0.00868, indicating that the effect is statistically significant at the 1% level ($p < 0.01$). This suggests that receiving ZMapp is associated with a lower likelihood of dehydration on day 14 compared to the control group.

4.2 Add baseline covariates

Add covariates for dehydration on day 0 and patient temperature on day 0 to the regression from part (a) and report the ATE (with standard error). Also report the p-value.

```r
zmapp_2 <- glm(dehydrated_day14 ~ treat_zmapp + dehydrated_day0 + temperature_day0, data = d, family = 
summary(zmapp_2)
```

```
##
## Call:
## glm(formula = dehydrated_day14 ~ treat_zmapp + dehydrated_day0 +
##     temperature_day0, family = binomial, data = d)
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -228.9320    81.6951  -2.802  0.00507 **
## treat_zmapp        -1.1979     0.5766  -2.078  0.03774 *
## dehydrated_day0    -1.0126     1.1522  -0.879  0.37946
## temperature_day0    2.3549     0.8388   2.807  0.00499 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 112.467  on 99  degrees of freedom
## Residual deviance:  75.987  on 96  degrees of freedom
## AIC: 83.987
##
## Number of Fisher Scoring iterations: 6
```

Answer: The estimated effect of ZMapp on dehydration on day 14, after adjusting for dehydration on day 0 and temperature on day 0, is -1.1979 with a standard error of 0.5766. The p-value for this estimate is 0.03774, indicating that the effect remains statistically significant at the 5% level ($p < 0.05$).

Including these baseline covariates has reduced the estimated effect slightly (from -1.2685 in the unadjusted model to -1.1979) and increased the standard error (from 0.4833 to 0.5766), but the effect remains statistically significant.

4.3 Interpret estimates

Do you prefer the estimate of the ATE reported in the chunk called `dehydration model` or `add pre-treatment measures`? Why? Report the results of the F-test that you used to form this opinion.

```r
zmapp_test_object <- anova(zmapp_1, zmapp_2, test = "Chisq")
zmapp_test_object
```

```
## Analysis of Deviance Table
##
## Model 1: dehydrated_day14 ~ treat_zmapp
## Model 2: dehydrated_day14 ~ treat_zmapp + dehydrated_day0 + temperature_day0
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
## 1        98    105.243
## 2        96     75.987  2   29.257 4.436e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer: I prefer the estimate reported in the add pre-treatment measures model (zmapp_2) because it accounts for baseline dehydration and temperature, which are important covariates that likely influence the outcome.

The Chi-square test result comparing the two models shows a highly significant improvement when adding these covariates ($^2(2) = 29.257$, p = 4.436e-07). This suggests that including dehydration on day 0 and temperature on day 0 significantly improves the model's explanatory power.

By controlling for these baseline characteristics, we obtain a more precise and adjusted estimate of the treatment effect, which helps isolate the true impact of ZMapp from pre-existing differences between patients.

4.4 Add day fourteen temperature

The regression from part `add pre-treatment measures` suggests that temperature is highly predictive of dehydration. Add, temperature on day 14 as a covariate and report the ATE, the standard error, and the p-value.

```r
zmapp_3 <- glm(dehydrated_day14 ~ treat_zmapp + dehydrated_day0 + temperature_day0 + temperature_day14,
summary(zmapp_3)
```

```
##
## Call:
```

```
## glm(formula = dehydrated_day14 ~ treat_zmapp + dehydrated_day0 +
##     temperature_day0 + temperature_day14, family = binomial,
##     data = d)
##
## Coefficients:
##                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -321.5889   100.4724  -3.201  0.00137 **
## treat_zmapp         -1.0196     0.6143  -1.660  0.09698 .
## dehydrated_day0     -1.5240     1.2474  -1.222  0.22179
## temperature_day0     2.2378     0.8936   2.504  0.01227 *
## temperature_day14    1.0586     0.4602   2.300  0.02143 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 112.467  on 99  degrees of freedom
## Residual deviance:  67.752  on 95  degrees of freedom
## AIC: 77.752
##
## Number of Fisher Scoring iterations: 6
```

Answer: After adding temperature on day 14 as an additional covariate, the estimated effect of ZMapp on dehydration remains negative but is less statistically significant than in previous models: - ATE (treat_zmapp estimate): -1.0196 - Standard Error: 0.6143 - p-value: 0.09698

This suggests that when accounting for temperature on day 14, the estimated effect of ZMapp on dehydration is still negative, but with a weaker statistical significance (p   0.097).

4.5 Interpret estimates

Do you prefer the estimate of the ATE reported in part `add pre-treatment measures` or `add day 14 temperature`? What is this preference based on?

Answer: I prefer the estimate from add pre-treatment measures (zmapp_2) because adding temperature on day 14 (zmapp_3) likely absorbs part of ZMapp's effect, making the treatment appear less significant (p-value increases from 0.038 to 0.097). Since temperature on day 14 is a post-treatment variable, controlling for it may underestimate ZMapp's true impact on dehydration.

4.6 Look at temperature

Now let's switch from the outcome of dehydration to the outcome of temperature, and use the same regression covariates as in the chunk titled `add pre-treatment measures`. Test the hypothesis that ZMapp is especially likely to reduce mens' temperatures, as compared to womens', and describe how you did so. What do the results suggest?

```r
zmapp_4 <- lm(temperature_day14 ~ treat_zmapp * male + temperature_day0 + dehydrated_day0, data = d)
summary(zmapp_4)
```

```
##
## Call:
## lm(formula = temperature_day14 ~ treat_zmapp * male + temperature_day0 +
##     dehydrated_day0, data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -0.70157 -0.37725 -0.02702  0.34687  0.73968
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       48.71269    9.26618   5.257 9.14e-07 ***
## treat_zmapp       -0.23087    0.11871  -1.945   0.0548 .
## male               3.08549    0.12644  24.403  < 2e-16 ***
## temperature_day0   0.50480    0.09508   5.309 7.34e-07 ***
## dehydrated_day0    0.04113    0.18208   0.226   0.8218
## treat_zmapp:male  -2.07669    0.19164 -10.836  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4518 on 94 degrees of freedom
## Multiple R-squared:  0.9059, Adjusted R-squared:  0.9009
## F-statistic:   181 on 5 and 94 DF,  p-value: < 2.2e-16
```

Answer: The interaction term (treat_zmapp:male) is highly significant ($p < 0.001$) and negative, suggesting that ZMapp reduces temperature more in men than in women. The coefficient of -2.08 indicates that while treatment alone has a small, marginally significant effect ($p = 0.055$), its effect varies substantially by gender.

Men have higher baseline temperatures (male coefficient = 3.09, $p < 0.001$), but ZMapp significantly lowers their temperature compared to women. This suggests a heterogeneous treatment effect, meaning the drug may work differently across genders.

4.7 Compare health outcomes

Which group – those that are coded as `male == 0` or `male == 1` have better health outcomes (temperature) in control? What about in treatment? How does this help to contextualize whatever heterogeneous treatment effect you might have estimated?

```
aggregate(temperature_day14 ~ male + treat_zmapp, data = d, mean)
```

```
##   male treat_zmapp temperature_day14
## 1    0           0          98.48654
## 2    1           0         101.69167
## 3    0           1          98.16308
## 4    1           1          99.10071
```

Answer: In the control group (treat_zmapp = 0), men (male = 1) have a higher average temperature (101.69) compared to women (98.49). However, in the treatment group (treat_zmapp = 1), men's average temperature decreases significantly to 99.10, while women's temperature remains relatively stable at 98.16.

This supports the earlier finding that ZMapp has a stronger temperature-reducing effect for men than for women, indicating a heterogeneous treatment effect by gender.

4.8 Collaborating with others, Part (1)

Suppose you speak with a colleague to learn about heterogeneous treatment effects.

This colleague has access to a non-anonymized version of the same dataset and reports that they looked at heterogeneous effects of the ZMapp treatment by each of 80 different covariates to examine whether each predicted the effectiveness of ZMapp on each of 20 different indicators of health.

Across these regressions your colleague ran, the treatment's interaction with sex on the outcome of temperature is the only heterogeneous treatment effect that he found to be statistically significant. They reason that this shows the importance of sex for understanding the effectiveness of the drug, because nothing else seemed to indicate why it worked. Bolstering your colleague's confidence, after looking at the data, they also returned to their medical textbooks see the whispers of a theory about why ZMapp interacts with processes only present in men to cure.

Another doctor, unfamiliar with the data, hears your colleague's theory and finds it plausible. How likely do you think it is ZMapp works especially well for curing Ebola in men, and why? (This question is conceptual can be answered without performing any computation.)

Answer: The likelihood that ZMapp works especially well for men is uncertain. Since the result was found after testing 80 covariates across 20 outcomes, it could be due to multiple comparisons rather than a true effect. Without proper correction for multiple testing, a statistically significant result could emerge by chance. The post-hoc search for a biological explanation also raises confirmation bias concerns. Further pre-registered studies or replications are needed before concluding that sex influences ZMapp's effectiveness.

4.9 Collaborating with others, Part (2)

Suppose that your colleague conducted their research looking at the interaction of 80 covariates with ZMapp, but that you on your own tested this and only this HTE, and discovered a positive result. How, if at all, does your colleague's behavior change the interpretation of your test? Does this seem fair or reasonable?

Answer: Since my colleague tested 80 covariates and only reported this one significant interaction, the risk of p-hacking or data dredging is high. This undermines confidence in the result because it may have been found by chance rather than reflecting a real effect. In contrast, if I had only tested this hypothesis independently, it would be less prone to such biases. Their broad search weakens the reliability of the finding and suggests the need for further validation.

4.10 Collaborating with others, Part (3)

Now, imagine that your colleague had not conducted the 80 different regressions. Instead, they tested this heterogeneous treatment effect, and only this heterogeneous treatment effect, of their own accord. Would you be more or less inclined to believe that the heterogeneous treatment effect really exists? Why? Is there a general principle that is guiding your reasoning?

Answer: If my colleague had tested only this one hypothesis from the start, I would be more inclined to believe the effect is real. A single pre-specified test reduces the risk of false positives caused by multiple comparisons. The general principle here is that selective reporting and post-hoc analysis inflate false discovery rates. Findings from exploratory analyses should be considered hypothesis-generating, requiring further confirmation in independent studies.