

# PSet 3

11/10/2018

ID 2174137

1. For this question, restrict your analysis to the set of job applications that asked about criminal records (“Box” applications) in the before period (“pre-BTB” period). (Note: there are some applications that did not have a box in the pre-BTB period, but then added them in the post- period. Agan and Star code these as “remover = -1” in their data and call them “reverse compliers.” Exclude these observations from your analysis throughout this assignment.)

a) What is the average callback rate for people who committed crimes? For those who didn’t? Is the difference statistically significant?

```
crime <- subset(newset, crime==1)
nocrime <- subset(newset, crime==0)
t.test(crime$response, nocrime$response)

##
## Welch Two Sample t-test
##
## data: crime$response and nocrime$response
## t = -4.2171, df = 2533.4, p-value = 2.561e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.07490189 -0.02735417
## sample estimates:
## mean of x mean of y
## 0.08458084 0.13570887
```

The average callback rate for people who committed crimes is 8.46%, while the average call back rate for those who did not commit a crime is 13.57%. Since the t-statistic is greater than 2, and the p-value is small, we can reject the null hypothesis of no difference between the means. Though we haven’t proven that they’re different, we can reject the idea that the means are not identical.

b) Can we interpret this as a causal effect? Explain briefly.

We cannot interpret interpret this as a causal effect given that in this analysis, we’re ignoring a whole host of other observable characteristics that might be relevant. There might be other characteristics in the data that are correlated with criminality, which means that in order to establish a causal argument, we need to analyze the other observable traits of the data.

2) Now consider just the “Box” applications but include both the pre- and post-BTB periods.

a) Regress callback rates on race, GED, and employment gap. Include “chain”1 and “center” fixed effects. Does race appear to have an effect on callback rates? Does this coefficient have a causal interpretation?

```
FEReg_white <- lm(response ~ white + ged + empgap + factor(cogroup_comb) + factor(center), data = q2set)
FEReg_black <- lm(response ~ black + ged + empgap + factor(cogroup_comb) + factor(center), data = q2set)
stargazer(FEReg_white, FEReg_black, type='text', title='Callback Rates on Race')
```

```
##
## Callback Rates on Race
## =====
##                               Dependent variable:
##                               -----
##                               response
##                               (1)         (2)
## -----
## white                        0.024***
##                               (0.009)
##
## black                        -0.024***
##                               (0.009)
##
## ged                          0.006
##                               (0.009)
##
## empgap                       0.006
##                               (0.009)
##
## factor(cogroup_comb)46       0.175**
##                               (0.085)
##
## factor(cogroup_comb)54       -0.044
##                               (0.096)
##
## factor(cogroup_comb)74       0.126*
##                               (0.071)
##
## factor(cogroup_comb)78       0.118
##                               (0.095)
##
## factor(cogroup_comb)101      -0.029
##                               (0.075)
##
## factor(cogroup_comb)102      0.076
##                               (0.097)
##
## factor(cogroup_comb)109      -0.041
##                               (0.089)
##
## factor(cogroup_comb)120      0.534***
##                               0.534***
```

##	(0.103)	(0.103)
##		
## factor(cogroup_comb)121	0.098	0.098
##	(0.078)	(0.078)
##		
## factor(cogroup_comb)154	0.190*	0.190*
##	(0.113)	(0.113)
##		
## factor(cogroup_comb)160	0.133	0.133
##	(0.085)	(0.085)
##		
## factor(cogroup_comb)161	0.035	0.035
##	(0.174)	(0.174)
##		
## factor(cogroup_comb)163	0.354***	0.354***
##	(0.081)	(0.081)
##		
## factor(cogroup_comb)168	0.008	0.008
##	(0.082)	(0.082)
##		
## factor(cogroup_comb)171	0.021	0.021
##	(0.071)	(0.071)
##		
## factor(cogroup_comb)179	-0.024	-0.024
##	(0.076)	(0.076)
##		
## factor(cogroup_comb)181	-0.037	-0.037
##	(0.139)	(0.139)
##		
## factor(cogroup_comb)185	0.160*	0.160*
##	(0.086)	(0.086)
##		
## factor(cogroup_comb)187	-0.105	-0.105
##	(0.143)	(0.143)
##		
## factor(cogroup_comb)194	0.227***	0.227***
##	(0.088)	(0.088)
##		
## factor(cogroup_comb)200	0.070	0.070
##	(0.074)	(0.074)
##		
## factor(cogroup_comb)203	0.082	0.082
##	(0.082)	(0.082)
##		
## factor(cogroup_comb)205	0.381***	0.381***
##	(0.094)	(0.094)
##		
## factor(cogroup_comb)209	-0.058	-0.058
##	(0.076)	(0.076)
##		
## factor(cogroup_comb)220	0.183	0.183
##	(0.147)	(0.147)
##		
## factor(cogroup_comb)225	0.472***	0.472***

##	(0.107)	(0.107)
##		
## factor(cogroup_comb)227	0.008	0.008
##	(0.089)	(0.089)
##		
## factor(cogroup_comb)230	0.146	0.146
##	(0.102)	(0.102)
##		
## factor(cogroup_comb)241	0.229***	0.229***
##	(0.074)	(0.074)
##		
## factor(cogroup_comb)246	-0.033	-0.033
##	(0.116)	(0.116)
##		
## factor(cogroup_comb)247	-0.042	-0.042
##	(0.102)	(0.102)
##		
## factor(cogroup_comb)252	-0.009	-0.009
##	(0.079)	(0.079)
##		
## factor(cogroup_comb)259	0.050	0.050
##	(0.116)	(0.116)
##		
## factor(cogroup_comb)289	0.188**	0.188**
##	(0.076)	(0.076)
##		
## factor(cogroup_comb)297	0.122	0.122
##	(0.091)	(0.091)
##		
## factor(cogroup_comb)90001	0.054	0.054
##	(0.134)	(0.134)
##		
## factor(cogroup_comb)90002	-0.091	-0.091
##	(0.098)	(0.098)
##		
## factor(cogroup_comb)90005	0.046	0.046
##	(0.113)	(0.113)
##		
## factor(cogroup_comb)90008	0.052	0.052
##	(0.093)	(0.093)
##		
## factor(cogroup_comb)90012	0.475***	0.475***
##	(0.134)	(0.134)
##		
## factor(cogroup_comb)90013	-0.059	-0.059
##	(0.147)	(0.147)
##		
## factor(cogroup_comb)90022	-0.022	-0.022
##	(0.073)	(0.073)
##		
## factor(cogroup_comb)90023	0.015	0.015
##	(0.082)	(0.082)
##		
## factor(cogroup_comb)90024	0.217**	0.217**

##	(0.090)	(0.090)
##		
## factor(cogroup_comb)90025	-0.030	-0.030
##	(0.127)	(0.127)
##		
## factor(cogroup_comb)90050	0.122	0.122
##	(0.117)	(0.117)
##		
## factor(center)2	0.024	0.024
##	(0.052)	(0.052)
##		
## factor(center)3	0.027	0.027
##	(0.055)	(0.055)
##		
## factor(center)4	0.017	0.017
##	(0.053)	(0.053)
##		
## factor(center)5	-0.073	-0.073
##	(0.056)	(0.056)
##		
## factor(center)6	-0.055	-0.055
##	(0.050)	(0.050)
##		
## factor(center)7	-0.072	-0.072
##	(0.053)	(0.053)
##		
## factor(center)8	-0.057	-0.057
##	(0.055)	(0.055)
##		
## factor(center)9	-0.025	-0.025
##	(0.053)	(0.053)
##		
## factor(center)10	-0.057	-0.057
##	(0.052)	(0.052)
##		
## factor(center)11	0.012	0.012
##	(0.052)	(0.052)
##		
## factor(center)12	-0.004	-0.004
##	(0.052)	(0.052)
##		
## factor(center)13	-0.019	-0.019
##	(0.057)	(0.057)
##		
## factor(center)14	-0.067	-0.067
##	(0.051)	(0.051)
##		
## factor(center)15	0.005	0.005
##	(0.049)	(0.049)
##		
## factor(center)16	0.004	0.004
##	(0.052)	(0.052)
##		
## factor(center)17	0.007	0.007

##	(0.051)	(0.051)
##		
## factor(center)18	-0.135	-0.135
##	(0.101)	(0.101)
##		
## factor(center)19	-0.035	-0.035
##	(0.053)	(0.053)
##		
## factor(center)20	-0.047	-0.047
##	(0.052)	(0.052)
##		
## factor(center)21	0.104*	0.104*
##	(0.061)	(0.061)
##		
## factor(center)22	-0.008	-0.008
##	(0.054)	(0.054)
##		
## factor(center)23	-0.016	-0.016
##	(0.052)	(0.052)
##		
## factor(center)24	-0.003	-0.003
##	(0.053)	(0.053)
##		
## factor(center)25	-0.051	-0.051
##	(0.063)	(0.063)
##		
## factor(center)26	0.072	0.072
##	(0.057)	(0.057)
##		
## factor(center)27	0.026	0.026
##	(0.049)	(0.049)
##		
## factor(center)28	0.050	0.050
##	(0.054)	(0.054)
##		
## factor(center)29	-0.008	-0.008
##	(0.050)	(0.050)
##		
## factor(center)30	-0.028	-0.028
##	(0.058)	(0.058)
##		
## factor(center)31	0.034	0.034
##	(0.056)	(0.056)
##		
## factor(center)32	0.038	0.038
##	(0.056)	(0.056)
##		
## factor(center)33	0.007	0.007
##	(0.056)	(0.056)
##		
## factor(center)34	-0.012	-0.012
##	(0.052)	(0.052)
##		
## factor(center)35	0.018	0.018

##	(0.052)	(0.052)
##		
## factor(center)36	0.010	0.010
##	(0.051)	(0.051)
##		
## factor(center)37	0.004	0.004
##	(0.056)	(0.056)
##		
## factor(center)38	0.091	0.091
##	(0.060)	(0.060)
##		
## factor(center)39	-0.022	-0.022
##	(0.059)	(0.059)
##		
## factor(center)40	0.023	0.023
##	(0.053)	(0.053)
##		
## factor(center)41	0.024	0.024
##	(0.057)	(0.057)
##		
## factor(center)42	-0.015	-0.015
##	(0.057)	(0.057)
##		
## factor(center)43	0.076	0.076
##	(0.051)	(0.051)
##		
## factor(center)44	0.033	0.033
##	(0.055)	(0.055)
##		
## factor(center)101	0.112*	0.112*
##	(0.058)	(0.058)
##		
## factor(center)102	0.140**	0.140**
##	(0.057)	(0.057)
##		
## factor(center)103	-0.154**	-0.154**
##	(0.069)	(0.069)
##		
## factor(center)104	0.014	0.014
##	(0.052)	(0.052)
##		
## factor(center)105	0.057	0.057
##	(0.052)	(0.052)
##		
## factor(center)106	0.045	0.045
##	(0.053)	(0.053)
##		
## factor(center)107	0.007	0.007
##	(0.051)	(0.051)
##		
## factor(center)108	0.039	0.039
##	(0.055)	(0.055)
##		
## factor(center)109	0.190***	0.190***

##	(0.064)	(0.064)
##		
## factor(center)110	-0.032	-0.032
##	(0.048)	(0.048)
##		
## factor(center)111	-0.045	-0.045
##	(0.046)	(0.046)
##		
## factor(center)112	0.113**	0.113**
##	(0.056)	(0.056)
##		
## factor(center)113	-0.099	-0.099
##	(0.075)	(0.075)
##		
## factor(center)114	0.121*	0.121*
##	(0.066)	(0.066)
##		
## factor(center)115	-0.045	-0.045
##	(0.079)	(0.079)
##		
## factor(center)116	0.043	0.043
##	(0.049)	(0.049)
##		
## factor(center)117	0.086	0.086
##	(0.055)	(0.055)
##		
## factor(center)118	0.092	0.092
##	(0.086)	(0.086)
##		
## factor(center)119	0.026	0.026
##	(0.072)	(0.072)
##		
## factor(center)120	0.052	0.052
##	(0.055)	(0.055)
##		
## factor(center)121	-0.063	-0.063
##	(0.057)	(0.057)
##		
## factor(center)122	-0.025	-0.025
##	(0.048)	(0.048)
##		
## factor(center)123	0.026	0.026
##	(0.097)	(0.097)
##		
## factor(center)124	0.003	0.003
##	(0.053)	(0.053)
##		
## factor(center)125	0.058	0.058
##	(0.061)	(0.061)
##		
## factor(center)126	0.231***	0.231***
##	(0.060)	(0.060)
##		
## factor(center)127	-0.010	-0.010



##	(0.073)	(0.073)
##		
## factor(center)128	0.024	0.024
##	(0.086)	(0.086)
##		
## factor(center)129	0.157***	0.157***
##	(0.052)	(0.052)
##		
## factor(center)130	0.052	0.052
##	(0.065)	(0.065)
##		
## factor(center)131	0.052	0.052
##	(0.056)	(0.056)
##		
## factor(center)132	0.104	0.104
##	(0.086)	(0.086)
##		
## factor(center)133	-0.148*	-0.148*
##	(0.078)	(0.078)
##		
## factor(center)134	0.021	0.021
##	(0.060)	(0.060)
##		
## factor(center)135	0.094*	0.094*
##	(0.051)	(0.051)
##		
## factor(center)136	0.078	0.078
##	(0.069)	(0.069)
##		
## factor(center)137	-0.075	-0.075
##	(0.056)	(0.056)
##		
## factor(center)138	-0.025	-0.025
##	(0.064)	(0.064)
##		
## factor(center)139	0.079	0.079
##	(0.074)	(0.074)
##		
## factor(center)140	0.113**	0.113**
##	(0.056)	(0.056)
##		
## Constant	0.018	0.042
##	(0.080)	(0.080)
##		
## -----		
## Observations	4,794	4,794
## R2	0.123	0.123
## Adjusted R2	0.099	0.099
## Residual Std. Error (df = 4661)	0.308	0.308
## F Statistic (df = 132; 4661)	4.975***	4.975***
## =====		
## Note:	*p<0.1; **p<0.05; ***p<0.01	

Given the tables above, race does appear to have an effect on callback rates when including fixed effects. For employers that asked for criminal records both before the BTB policy change but not after (since we subsetting out non-compliers), being white is associated with a 2.4% increase on callbacks, while being black is associated with a 2.4% decrease on callbacks. Both effects are statistically significant, while GED and employment gap are not, thus indicating that race could be a causal mechanism for callback rates for these companies.

b) Estimate the model again, but without the chain and center fixed effects. Does the coefficient on “white” change? Why is it important to include chain and center fixed effects?

```
reg.white <- lm(response ~ white + ged + empgap, data = q2set)
reg.black <- lm(response ~ black + ged + empgap, data = q2set)
stargazer(reg.white, reg.black, type='text', title='Callback Rates on Race w/o FE')
```

```
##
## Callback Rates on Race w/o FE
## =====
##                               Dependent variable:
##                               -----
##                               response
##                               (1)         (2)
## -----
## white                        0.024***
##                               (0.009)
##
## black                        -0.024***
##                               (0.009)
##
## ged                          0.008      0.008
##                               (0.009)    (0.009)
##
## empgap                       0.006      0.006
##                               (0.009)    (0.009)
##
## Constant                     0.101***    0.125***
##                               (0.009)    (0.009)
## -----
## Observations                 4,794      4,794
## R2                           0.002      0.002
## Adjusted R2                  0.001      0.001
## Residual Std. Error (df = 4790) 0.325      0.325
## F Statistic (df = 3; 4790)      2.579*      2.579*
## =====
## Note:                         *p<0.1; **p<0.05; ***p<0.01
```

Generally, including fixed effects reduces bias by controlling, in this case, for the center that a person is applying from and the chain store they are applying to. The coefficient on white changes very slightly, this indicates that the center that a white person is applying from has no significant bearing on the callback rates as it relates to race, and neither does the chain store that a white person is applying to. In so many words, being white is associated with the same rate of callbacks independent of which center they apply from, or which chain store

they apply to.

c) Now add the “conviction” variable. What happens to the coefficient on “white”? If the coefficient changes, does this mean that the previous regression was subject to omitted variable bias?

```
reg.whiteomv <- lm(response ~ white + ged + empgap + crime, data = q2set)
stargazer(reg.whiteomv, type='text', title='Callback Rates on Race w/Conviction')
```

```
##
## Callback Rates on Race w/Conviction
## =====
##                               Dependent variable:
##                               -----
##                               response
## -----
## white                        0.025***
##                               (0.009)
##
## ged                          0.007
##                               (0.009)
##
## empgap                       0.006
##                               (0.009)
##
## crime                        -0.037***
##                               (0.009)
##
## Constant                     0.120***
##                               (0.011)
##
## -----
## Observations                 4,794
## R2                           0.005
## Adjusted R2                  0.004
## Residual Std. Error         0.324 (df = 4789)
## F Statistic                  5.846*** (df = 4; 4789)
## =====
## Note:                        *p<0.1; **p<0.05; ***p<0.01
```

The coefficient on white doesn’t change that much, it shifts from 2.4% to 2.5%. Though the change is small, it is an increase in callback rates for white men once you control for conviction. This indicates that there might be a slight omitted variable bias effect – as you control for the criminal history of an applicant, the callback rate goes up slightly, which is a result we would expect.

3) The authors estimate the following model for different subsets of the data, where “Box” is an indicator for whether the application had a box asking about employment<sup>2</sup>, and  $X$  is a vector of covariates:

a) Suppose they run this regression on the full sample, which includes both Box and non-Box applications, but only in the pre-period (don’t actually do this yet). What do  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  tell you?

If an employer includes a box on an application prior to the ‘Ban the Box’ initiative, the variables mean the following:

1) Intercept value is the expected mean value of callbacks prior to the policy change when an individual is not white and the application does not have a box.

2) The beta one value indicates the size of the pre-policy effect on callback rates if the application has a box.

3) The beta two value indicates the size of the pre-policy effect on callback rates if the applicant is white

4) The beta three value indicates the size of the pre-policy effect on callback rates if the applicant is white and there is a box on the application.

b) Do you think “Box” and “non-Box” stores might differ in systematic ways, besides their decision to include a box asking about criminal history? In other words, do we think this variable is “as-if” randomly assigned?

Yes, I think Box and non-Box stores differ in systematic ways. For example, I think a Box Store is more likely to be a retail store that wants to make sure it’s protecting inventory by not hiring individuals with criminal records, particularly theft-related histories; meanwhile a non-box store is more likely to be a food-service job that tends to have significant turnover and where perceived criminal proclivities might not be as damaging to the business. Due to those differences, it could be the case that a Box store is located in more upscale neighborhoods where it might behoove them to include the box, while a large, non-Box store could be located in tougher neighborhoods where including a box could limit the pool of applicants.

c) Suppose they run the regression on just the “Box” applications in both periods (again, don’t do this yet). What is the interpretation of the coefficients now?

If the regression is run on a subset of the sample that’s just application with the box, this means that we’re only taking into account businesses that actively want to have the box in the pre-period, but that don’t have it in the post-period (since we subsetting out non-compliers). In that case, the variables mean the following:

1) Intercept value is the expected mean value of callbacks for non-white applicants.

2) The beta one variable is the difference in callback rates associated with companies that had a box in the pre-period and those that did not in the post-period.

3) The beta two variable in this scenario indicates size of the effect on callback rates if the applicant is white.

4) The beta three variable is the size of the effect on callback rates for a white applicant given an application with a box.

4) For the below estimations, include controls for employment gap and ged, as well as center fixed effects. Again, exclude the so-called “reverse compliers.”

a) Estimate the model from question 3 on both “Box” and non-“Box” applications in just the pre-period.

```
regr.pre <- lm(response ~ crimbox + white + ged + empgap + (crimbox * white) + factor(center), data = p
summary(regr.pre, cluster=c("chain_id"))
```

```
##
## Call:
## lm(formula = response ~ crimbox + white + ged + empgap + (crimbox *
##      white) + factor(center), data = preset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29911 -0.12801 -0.08745 -0.05326  0.99540
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0847219  0.0312519   2.711 0.006725 **
## crimbox        0.0150434  0.0107543   1.399 0.161907
## white          0.0315233  0.0091238   3.455 0.000553 ***
## ged           -0.0059729  0.0073476  -0.813 0.416299
## empgap        -0.0022208  0.0072975  -0.304 0.760894
## factor(center)2  0.0036386  0.0406491   0.090 0.928677
## factor(center)3  0.0032120  0.0459230   0.070 0.944241
## factor(center)4 -0.0034353  0.0424377  -0.081 0.935485
## factor(center)5 -0.0396030  0.0447389  -0.885 0.376077
## factor(center)6  0.0200292  0.0406993   0.492 0.622646
## factor(center)7 -0.0407470  0.0443412  -0.919 0.358157
## factor(center)8 -0.0349564  0.0434923  -0.804 0.421575
## factor(center)9 -0.0355148  0.0433711  -0.819 0.412894
## factor(center)10 -0.0336695  0.0421520  -0.799 0.424454
## factor(center)11  0.0405205  0.0400512   1.012 0.311707
## factor(center)12 -0.0124642  0.0406399  -0.307 0.759082
## factor(center)13  0.0304631  0.0460955   0.661 0.508717
## factor(center)14 -0.0201565  0.0415417  -0.485 0.627542
## factor(center)15 -0.0219883  0.0403670  -0.545 0.585970
## factor(center)16 -0.0358298  0.0433697  -0.826 0.408747
## factor(center)17 -0.0044646  0.0434729  -0.103 0.918205
## factor(center)18 -0.0047385  0.0724696  -0.065 0.947868
## factor(center)19 -0.0421670  0.0421536  -1.000 0.317190
## factor(center)20 -0.0424894  0.0421495  -1.008 0.313458
## factor(center)21  0.0022010  0.0458034   0.048 0.961675
## factor(center)22 -0.0342362  0.0436097  -0.785 0.432444
```

```

## factor(center)23 -0.0184154 0.0417141 -0.441 0.658888
## factor(center)24 -0.0465017 0.0415433 -1.119 0.263026
## factor(center)25 -0.0165363 0.0495561 -0.334 0.738625
## factor(center)26 0.0319569 0.0477428 0.669 0.503291
## factor(center)27 0.0397805 0.0416230 0.956 0.339240
## factor(center)28 0.0435399 0.0432650 1.006 0.314279
## factor(center)29 -0.0599346 0.0423287 -1.416 0.156838
## factor(center)30 -0.0801208 0.0492989 -1.625 0.104163
## factor(center)31 -0.0153692 0.0422635 -0.364 0.716129
## factor(center)32 -0.0069761 0.0444793 -0.157 0.875376
## factor(center)33 0.0136724 0.0432649 0.316 0.752000
## factor(center)34 -0.0635960 0.0413066 -1.540 0.123700
## factor(center)35 -0.0100625 0.0428335 -0.235 0.814276
## factor(center)36 -0.0265078 0.0431449 -0.614 0.538977
## factor(center)37 0.0240788 0.0454615 0.530 0.596369
## factor(center)38 -0.0005017 0.0471502 -0.011 0.991511
## factor(center)39 -0.0549032 0.0435134 -1.262 0.207078
## factor(center)40 0.0365047 0.0454463 0.803 0.421857
## factor(center)41 -0.0353136 0.0435135 -0.812 0.417074
## factor(center)42 -0.0149825 0.0450218 -0.333 0.739308
## factor(center)43 0.0425814 0.0424421 1.003 0.315759
## factor(center)44 0.0086985 0.0419081 0.208 0.835578
## factor(center)101 0.1300246 0.0426590 3.048 0.002312 **
## factor(center)102 0.1186962 0.0591254 2.008 0.044731 *
## factor(center)103 0.0538031 0.0616987 0.872 0.383221
## factor(center)104 0.0684148 0.0443332 1.543 0.122828
## factor(center)105 0.0994582 0.0441985 2.250 0.024463 *
## factor(center)106 0.0895201 0.0435940 2.053 0.040061 *
## factor(center)107 0.0141413 0.0416521 0.340 0.734236
## factor(center)108 0.0612903 0.0465990 1.315 0.188461
## factor(center)109 0.1374370 0.0464357 2.960 0.003089 **
## factor(center)110 -0.0644853 0.0409301 -1.575 0.115185
## factor(center)111 0.0177132 0.0375613 0.472 0.637241
## factor(center)112 0.0813709 0.0453370 1.795 0.072727 .
## factor(center)113 0.1549499 0.0524620 2.954 0.003151 **
## factor(center)114 -0.0101531 0.0477740 -0.213 0.831705
## factor(center)115 -0.0572089 0.0677108 -0.845 0.398194
## factor(center)116 0.0607953 0.0439569 1.383 0.166687
## factor(center)117 0.1017590 0.0413336 2.462 0.013844 *
## factor(center)118 0.0306603 0.0699460 0.438 0.661152
## factor(center)119 0.0423626 0.0560355 0.756 0.449676
## factor(center)120 0.0758597 0.0421825 1.798 0.072161 .
## factor(center)121 -0.0534643 0.0482377 -1.108 0.267747
## factor(center)122 0.0820829 0.0399773 2.053 0.040086 *
## factor(center)123 0.0945805 0.0677370 1.396 0.162671
## factor(center)124 0.0460802 0.0490576 0.939 0.347603
## factor(center)125 0.0626930 0.0527869 1.188 0.235006
## factor(center)126 0.0236940 0.0467884 0.506 0.612587
## factor(center)127 -0.0244680 0.0560074 -0.437 0.662219
## factor(center)128 0.0042387 0.0753938 0.056 0.955167
## factor(center)129 0.1759401 0.0442068 3.980 6.96e-05 ***
## factor(center)130 -0.0195169 0.0524745 -0.372 0.709956
## factor(center)131 0.1690021 0.0468528 3.607 0.000312 ***
## factor(center)132 0.1850894 0.0688157 2.690 0.007170 **

```

```
## factor(center)133 -0.0338125  0.0624235  -0.542  0.588068
## factor(center)134 -0.0075125  0.0459800  -0.163  0.870219
## factor(center)135 -0.0067283  0.0411596  -0.163  0.870155
## factor(center)136  0.0422021  0.0648546   0.651  0.515249
## factor(center)137  0.0256326  0.0535108   0.479  0.631941
## factor(center)138 -0.0411874  0.0471895  -0.873  0.382797
## factor(center)139  0.1376850  0.0808000   1.704  0.088420 .
## factor(center)140 -0.0144945  0.0500922  -0.289  0.772317
## crimbox:white      -0.0302277  0.0150717  -2.006  0.044938 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3087 on 7156 degrees of freedom
## Multiple R-squared:  0.03408,    Adjusted R-squared:  0.0222
## F-statistic: 2.869 on 88 and 7156 DF,  p-value: < 2.2e-16
```

b) What kind of standard errors should you use, and why?

You should use chain-clustered standard errors, since the chain is likely correlated with the chance of getting a callback. Indeed, since chains are the units by which applications were clustered, it's likely the errors in those groups are correlated with each other. Not clustering might lead to underestimating our errors.

c) Is the coefficient on “crimbox” statistically significant? What about “white” and the interaction of “crimbox” and “white”? Interpret these findings.

The coefficient on crimbox is not statistically significant, while the coefficient on white is statistically significant. The interaction coefficient is similarly statistically significant but at a lesser rate ( $p < 0.001$  vs  $p < 0.05$ ). Essentially, this means that in the period prior to Ban the Box, the fact that certain applications had a box did not have a significant effect in callback rates, but that being white improved callback rates by 3.1%. The interaction coefficient, which doesn't meet criteria for statistical significance at  $p < 0.001$  but it does at  $p < 0.01$ , hints that a white candidate would have a lower callback rate if the application had a box.

d) Now estimate the model from question 3 on just “Box” applications in both periods. Interpret the coefficients.

```
regr.box <- lm(response ~ crimbox + white + ged + empgap + (crimbox * white) + factor(center), data = boxset)
summary(regr.box, cluster=c("chain_id"))

##
## Call:
## lm(formula = response ~ crimbox + white + ged + empgap + (crimbox *
##      white) + factor(center), data = boxset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38308 -0.14633 -0.09351 -0.05054  0.99528
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.129755   0.038130   3.403 0.000672 ***
```

```

## crimbox      -0.001747  0.013188 -0.133 0.894588
## white        0.040122  0.012991  3.089 0.002023 **
## ged          0.006882  0.009482  0.726 0.468031
## empgap       0.003483  0.009318  0.374 0.708591
## factor(center)2 -0.047746  0.052931 -0.902 0.367082
## factor(center)3  0.021223  0.057058  0.372 0.709944
## factor(center)4 -0.070825  0.053712 -1.319 0.187368
## factor(center)5 -0.127122  0.058032 -2.191 0.028533 *
## factor(center)6 -0.092959  0.051654 -1.800 0.071979 .
## factor(center)7 -0.101335  0.053746 -1.885 0.059429 .
## factor(center)8 -0.111617  0.055900 -1.997 0.045913 *
## factor(center)9 -0.073724  0.053306 -1.383 0.166717
## factor(center)10 -0.131920  0.053110 -2.484 0.013029 *
## factor(center)11 -0.032509  0.052916 -0.614 0.539015
## factor(center)12 -0.071850  0.053518 -1.343 0.179484
## factor(center)13 -0.084957  0.058355 -1.456 0.145498
## factor(center)14 -0.096007  0.051345 -1.870 0.061568 .
## factor(center)15 -0.052527  0.050459 -1.041 0.297934
## factor(center)16 -0.076307  0.052911 -1.442 0.149315
## factor(center)17 -0.024150  0.052338 -0.461 0.644518
## factor(center)18 -0.147410  0.103393 -1.426 0.154013
## factor(center)19 -0.084352  0.053925 -1.564 0.117826
## factor(center)20 -0.089084  0.052708 -1.690 0.091067 .
## factor(center)21  0.029917  0.062616  0.478 0.632826
## factor(center)22 -0.081576  0.054849 -1.487 0.137006
## factor(center)23 -0.088709  0.052900 -1.677 0.093624 .
## factor(center)24 -0.057447  0.053291 -1.078 0.281101
## factor(center)25 -0.067054  0.063709 -1.053 0.292621
## factor(center)26  0.010944  0.058012  0.189 0.850370
## factor(center)27  0.012753  0.050062  0.255 0.798925
## factor(center)28  0.018372  0.055116  0.333 0.738894
## factor(center)29 -0.060913  0.050747 -1.200 0.230076
## factor(center)30 -0.081963  0.059495 -1.378 0.168374
## factor(center)31  0.001319  0.057084  0.023 0.981564
## factor(center)32 -0.012503  0.057692 -0.217 0.828431
## factor(center)33 -0.021179  0.056445 -0.375 0.707515
## factor(center)34 -0.089077  0.052899 -1.684 0.092265 .
## factor(center)35 -0.026401  0.053495 -0.494 0.621666
## factor(center)36 -0.068883  0.051855 -1.328 0.184112
## factor(center)37 -0.052722  0.057042 -0.924 0.355389
## factor(center)38  0.092216  0.061631  1.496 0.134652
## factor(center)39 -0.101177  0.060292 -1.678 0.093387 .
## factor(center)40 -0.011101  0.053707 -0.207 0.836252
## factor(center)41 -0.043006  0.058714 -0.732 0.463919
## factor(center)42 -0.067115  0.058011 -1.157 0.247355
## factor(center)43  0.011676  0.052000  0.225 0.822343
## factor(center)44 -0.002398  0.056448 -0.042 0.966116
## factor(center)101 0.056284  0.058359  0.964 0.334872
## factor(center)102 0.109178  0.058043  1.881 0.060033 .
## factor(center)103 -0.145621  0.069957 -2.082 0.037433 *
## factor(center)104 -0.039732  0.052006 -0.764 0.444912
## factor(center)105  0.039480  0.052218  0.756 0.449652
## factor(center)106 -0.013354  0.053503 -0.250 0.802919
## factor(center)107 -0.017806  0.051869 -0.343 0.731399

```



```

## factor(center)108 0.022207 0.055667 0.399 0.689964
## factor(center)109 0.161529 0.063164 2.557 0.010580 *
## factor(center)110 -0.074752 0.047925 -1.560 0.118882
## factor(center)111 -0.059399 0.045624 -1.302 0.193014
## factor(center)112 0.094208 0.057081 1.650 0.098925 .
## factor(center)113 -0.025653 0.073895 -0.347 0.728492
## factor(center)114 0.094612 0.066821 1.416 0.156873
## factor(center)115 -0.093399 0.080564 -1.159 0.246389
## factor(center)116 -0.015343 0.048718 -0.315 0.752824
## factor(center)117 0.048752 0.055140 0.884 0.376662
## factor(center)118 0.030167 0.086012 0.351 0.725807
## factor(center)119 -0.029361 0.073938 -0.397 0.691306
## factor(center)120 0.037743 0.055387 0.681 0.495627
## factor(center)121 -0.072079 0.057721 -1.249 0.211820
## factor(center)122 -0.040353 0.048368 -0.834 0.404166
## factor(center)123 -0.062606 0.099551 -0.629 0.529457
## factor(center)124 -0.009470 0.054187 -0.175 0.861275
## factor(center)125 0.053376 0.060288 0.885 0.376012
## factor(center)126 0.202840 0.061159 3.317 0.000918 ***
## factor(center)127 -0.109875 0.074032 -1.484 0.137833
## factor(center)128 0.039410 0.088232 0.447 0.655136
## factor(center)129 0.122643 0.052370 2.342 0.019229 *
## factor(center)130 -0.001747 0.066150 -0.026 0.978929
## factor(center)131 -0.020203 0.056759 -0.356 0.721897
## factor(center)132 0.102625 0.088167 1.164 0.244491
## factor(center)133 -0.145011 0.079018 -1.835 0.066546 .
## factor(center)134 0.013035 0.060719 0.215 0.830031
## factor(center)135 0.076149 0.052004 1.464 0.143180
## factor(center)136 0.021306 0.070001 0.304 0.760866
## factor(center)137 -0.074550 0.056753 -1.314 0.189054
## factor(center)138 -0.069874 0.063147 -1.107 0.268555
## factor(center)139 0.066097 0.076329 0.866 0.386557
## factor(center)140 0.068814 0.056444 1.219 0.222840
## crimbox:white -0.032743 0.018549 -1.765 0.077591 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3205 on 4705 degrees of freedom
## Multiple R-squared: 0.04369, Adjusted R-squared: 0.0258
## F-statistic: 2.443 on 88 and 4705 DF, p-value: 3.019e-12

```

There only statistically significant result in this regression is the coefficient for white, which indicates that being white is associated with a 4% increase in callback rates for companies with the box in the pre-period, and without the box in the post-period.

## 5. Based on the above analysis, what are your conclusions about the effects of BTB?

From the results above, it appears that ban-the-box did not have the desired effect in helping reduce unemployment among black men, who comprise a disproportionate amount of individuals in the criminal justice system. In fact, from the analysis above it appears that being black is associated with lower callback rates in this data (see 2a), while being white is asso-

ciated with increases in callback rates in the post BTB period. This seems to indicate that post-BTB, employers relied on preconceived notions about race and criminality in order to make assumptions about the criminal histories of the applicants.

Mario Moreno