

(a) This part consists of seven questions:

- (i) The unit of observation in the dataset is “job application”.
- (ii) There are 14,637 observations in the dataset.
- (iii) The variable “*posresponse*” records if an applicant received a positive response; the variable “*interview*” records if an applicant got an interview request.
- (iv) Use conditional subsetting command of “count if (interview==1) & (posresponse==1)”, get 928 eligible applications. Therefore, the proportion of applications receiving a positive response and an interview request is calculated as  $928 \div 14,637 = 6.34\%$ .
- (v) The variable “*crime*” records if an applicant has criminal record; the variable *white* records if the applicant is white; the variable “*empgap*” records if the applicant has employment gap; the variable “*ged*” records if the applicant has GED.
- (vi) The variables “*pre*” (or “*post*”) records if an application took place before (or after) the ban.
- (vii) The variable “*box*” records if the job application asked about the criminal record of the applicant.

(b) Use “*table*” command to produce table of summary statistics in Table 1 as in the paper.

Due to formatting difficulties, the replicated tables below are transpose of original one.

(i) The first table corresponds to the “characteristics” panel:

```
. table post, contents(mean white mean crime mean ged mean empgap mean box)
```

| =1 if<br>post-BTB<br>period,<br>=0 if<br>pre-BTB<br>period |             |             |           |              |           |
|--|-------------|-------------|-----------|--------------|-----------|
|  | mean(white) | mean(crime) | mean(ged) | mean(empgap) | mean(box) |
| Pre  | .5024155    | .4970324    | .4977226  | .4923396     | .3664596  |
| Post   | .4967532    | .5128517    | .5021645  | .5041937     | .035579   |

(ii) The second table corresponds to the “results” panel:

```
. table post, contents(mean posresponse mean interview)
```

| =1 if<br>post-BTB<br>period,<br>=0 if<br>pre-BTB<br>period | mean(posres~e) mean(interv~w) |                 |
|--|-------------------------------|-----------------|
|  |                               |                 |
| Pre  | <b>.1094548</b>               | <b>.0597654</b> |
| Post   | <b>.1247294</b>               | <b>.0669643</b> |

(c) We want to determine if there is statistical significance to reject the null hypothesis, and below is the regression result of each observable characteristic on “*pre*”.

There are five observable characteristics that we regressed on. Observing the regression results as below, we can see all observables characteristics, except *box*, have a small t-statistic that fails to reject the null. The regression result of *box* gives a t-stat of 54.62, which is significant enough to reject the null, meaning this characteristic is statistically different before and after the ban. The results for *box* can be interpreted that on average 33.08 percentage points fewer employers asked about the criminal record of applicants Post-BTB than Pre-BTB. All other characteristics *white*, *crime*, *ged*, and *empgap* are not statistically different before and after the ban.

The research aimed to investigate the racial discrimination in job market in the absence of criminal declaration boxes, and the above findings comply to the research. The change in *box* can be explained by the enact of the BTB movement, which affected employers. No significant differences found in all other characteristics, since these are the control variables intentionally controlled by the investigators in fictional applications. It is important to have a difference in *box* with all other characteristics maintained. In that condition, we can conclude the job applications are randomly assigned and we can safely draw causality conclusions on the “removal of box” or BTB, with all lurking variables properly controlled.

```
. reg white pre, robust
```

```
Linear regression               Number of obs   =    14,637
                                F(1, 14635)      =      0.47
                                Prob > F          =    0.4934
                                R-squared         =    0.0000
                                Root MSE      =    .50003
```

|       | Coef.    | Robust<br>Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|-------|----------|---------------------|-------|-------|----------------------|----------|
| white |          |                     |       |       |                      |          |
| pre   | .0056622 | .0082664            | 0.68  | 0.493 | -.010541             | .0218655 |
| _cons | .4967532 | .0058158            | 85.41 | 0.000 | .4853535             | .508153  |

```
. reg crime pre, robust
```

```
Linear regression               Number of obs   =    14,637
                                F(1, 14635)      =      3.66
                                Prob > F          =    0.0556
                                R-squared         =    0.0003
                                Root MSE      =    .49995
```

|       | Coef.     | Robust<br>Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|-------|-----------|---------------------|-------|-------|----------------------|----------|
| crime |           |                     |       |       |                      |          |
| pre   | -.0158193 | .0082651            | -1.91 | 0.056 | -.03202              | .0003814 |
| _cons | .5128517  | .005814             | 88.21 | 0.000 | .5014556             | .5242479 |

```
. reg ged pre, robust
```

```
Linear regression               Number of obs   =    14,637
                                F(1, 14635)      =      0.29
                                Prob > F          =    0.5910
                                R-squared         =    0.0000
                                Root MSE      =    .50003
```

|       | Coef.     | Robust<br>Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|-------|-----------|---------------------|-------|-------|----------------------|----------|
| ged   |           |                     |       |       |                      |          |
| pre   | -.0044419 | .0082665            | -0.54 | 0.591 | -.0206453            | .0117614 |
| _cons | .5021645  | .0058159            | 86.34 | 0.000 | .4907647             | .5135643 |

```
. reg empgap pre, robust
```

```
Linear regression               Number of obs   =    14,637
                                F(1, 14635)      =      2.06
                                Prob > F          =    0.1516
                                R-squared         =    0.0001
                                Root MSE      =    .5
```

|        | Coef.     | Robust<br>Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|--------|-----------|---------------------|-------|-------|----------------------|----------|
| empgap |           |                     |       |       |                      |          |
| pre    | -.0118542 | .0082659            | -1.43 | 0.152 | -.0280565            | .0043481 |
| _cons  | .5041937  | .0058157            | 86.69 | 0.000 | .4927942             | .5155933 |

```
. reg box pre, robust
```

```
Linear regression              Number of obs   =    14,637
                               F(1, 14635)      =   2983.80
                               Prob > F          =    0.0000
                               R-squared         =    0.1715
                               Root MSE      =    .36368
```

| box   | Coef.    | Robust<br>Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|-------|----------|---------------------|-------|-------|----------------------|----------|
| pre   | .3308806 | .0060574            | 54.62 | 0.000 | .3190074             | .3427539 |
| _cons | .035579  | .0021547            | 16.51 | 0.000 | .0313556             | .0398024 |

(d) Summary for column (1):

White applicants were about 2.4 percentage points (23%) more likely to receive a callback after controlling for crime status, GED, employment gap, and time; after controlling for race, GED, employment gap, and time, about 1.4 percentage points fewer of those who with criminal records receive callback than those who has no criminal records.

Summary for column (2):

Limited to employers with the box, on average 5.2 percentage points more of applicant with criminal record received callback than those who has no criminal record; while no statistical difference was found in the characteristic of race.

|        | (1)                    | (2)                    |
|--------|------------------------|------------------------|
|        | posresponse            | posresponse            |
| white  | 0.0242***<br>(0.00563) | -0.000868<br>(0.00937) |
| crime  | -0.0129*<br>(0.00545)  | -0.0520***<br>(0.0123) |
| ged    | -0.00519<br>(0.00504)  | 0.0105<br>(0.0134)     |
| empgap | 0.00125<br>(0.00458)   | 0.0103<br>(0.0100)     |
| pre    | -0.0151<br>(0.0102)    |                        |
| N      | 14637                  | 2918                   |

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

(e) Similar with question (d), we generate the following output:

|        | (1)<br>interview       | (2)<br>interview        |
|--------|------------------------|-------------------------|
| white  | 0.0135**<br>(0.00471)  | -0.00740<br>(0.00633)   |
| crime  | -0.00931*<br>(0.00458) | -0.0353***<br>(0.00625) |
| ged    | -0.00239<br>(0.00339)  | -0.0000279<br>(0.00837) |
| empgap | -0.00305<br>(0.00287)  | 0.00662<br>(0.00720)    |
| pre    | -0.00715<br>(0.00918)  |                         |
| N      | 14637                  | 2918                    |

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

(f) The research is interested in the effect of BTB program, so all other observable confounders should be properly controlled. As mentioned in the paper, the opening and closure of stores between the two time periods, Pre-BTB and Post-BTB, were not randomized and can be of concern. Thus, this factor is a confounder in our analysis of the relationship between callback rate and the BTB movement. The fact that the stores were not identical across periods is problematic, since it added endogeneity bias and hinder us from drawing causality conclusions. In the paper, “59% of sample was sent only to one store”, and the investigations were furthered by handing out same pairs of resumes to nearby stores of the same chain, to minimize the confounding effect. Also, the paper mentioned the research in NYC, they limited the Post-BTB wave to stores that they had sent at least one application to Pre-BTB.

(g) The variable “balanced” denotes if an application is sent to a store present in both time periods, Pre-BTB and Post-BTB. Exactly four applications are sent to such stores, including one white/black application pair in each time period. As summarized below, 11, 188 applications were sent out to such stores.

. tab balanced

| Application<br>is in<br>balanced<br>sample | Freq.  | Percent | Cum.   |
|--|--------|---------|--------|
| 0  | 3,449  | 23.56   | 23.56  |
| 1  | 11,188 | 76.44   | 100.00 |
| Total                                      | 14,637 | 100.00  |        |

(h) To begin with, generate interaction terms *boxwhite* and *prewhite* to prepare for the regressions. For all columns, we controlled for *ged* and *empgap* as specified in the notes below Table IV. In column (1), we controlled for *center*; in column (2), we only included balanced observations that had the box and removed it after BTB; in column (3), we controlled for *center* and included only the observations with the box pre-BTB and removed it post-BTB; in column (4), we only included the balanced samples with the box pre-BTB and did not remove the box post-BTB.

|          | (1)<br>posresponse   | (2)<br>posresponse    | (3)<br>posresponse   | (4)<br>posresponse   |
|----------|----------------------|-----------------------|----------------------|----------------------|
| boxwhite | -0.0302*<br>(0.0153) | -0.0364*<br>(0.0139)  | -0.0327*<br>(0.0138) |                      |
| white    | 0.0315**<br>(0.0120) | 0.0442***<br>(0.0126) | 0.0401**<br>(0.0117) | 0.0221*<br>(0.00882) |
| box      | 0.0150<br>(0.0245)   | 0.00317<br>(0.0149)   | -0.00175<br>(0.0133) |                      |
| prewhite |                      |                       |                      | 0.00203<br>(0.0138)  |
| pre      |                      |                       |                      | -0.0159<br>(0.0175)  |
| N        | 7245                 | 3712                  | 4794                 | 7476                 |

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

(Note: in this table, column (4) corresponds to the column (5) in the original paper.)

(i) Similar with question (h), we generate the following output:

|          | (1)<br>interview      | (2)<br>interview     | (3)<br>interview      | (4)<br>interview      |
|----------|-----------------------|----------------------|-----------------------|-----------------------|
| boxwhite | -0.0262**<br>(0.0101) | -0.0409*<br>(0.0163) | -0.0360**<br>(0.0132) |                       |
| white    | 0.0169*<br>(0.00701)  | 0.0399*<br>(0.0181)  | 0.0323*<br>(0.0144)   | 0.0158*<br>(0.00680)  |
| box      | 0.0226<br>(0.0155)    | 0.0108<br>(0.00904)  | 0.00563<br>(0.00828)  |                       |
| prewhite |                       |                      |                       | -0.00237<br>(0.00865) |
| pre      |                       |                      |                       | -0.00580<br>(0.0155)  |
| N        | 7245                  | 3712                 | 4794                  | 7476                  |

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

(j) To investigate the similarity between the two variables *callback* and *interview*, we want to check their conclusion on statistical significance in both cases. As we generated two pairs of tables above, we will focus on comparisons between table (d) & (e) and between table (h) & (i), to test if the two variables have same results.

(i) Table (d) & (e)

The two tables are re-run and attached below, with the t-stat for each estimate shown.

Under observations with the 5% significance level, we can see all corresponding estimates remain consistent and share similar statistical significance conclusions across variables *callback* and *interview*. The results are quite similar, white applicants were on average more likely to receive a callback (or interview invite), but the effect is minimal if there is a box for criminal record in resume; also, applicants with criminal records generally receive fewer callback (or interview invite).

```
. esttab model_3 model_4, drop(_cons *.center *.chain_id) . esttab model_1 model_2, drop(_cons *.center *.chain_id)
```

|        | (1)<br>interview     | (2)<br>interview      |
|--------|----------------------|-----------------------|
| white  | 0.0135**<br>(2.87)   | -0.00740<br>(-1.17)   |
| crime  | -0.00931*<br>(-2.03) | -0.0353***<br>(-5.64) |
| ged    | -0.00239<br>(-0.71)  | -0.0000279<br>(-0.00) |
| empgap | -0.00305<br>(-1.06)  | 0.00662<br>(0.92)     |
| pre    | -0.00715<br>(-0.78)  |                       |
| N      | 14637                | 2918                  |

t statistics in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

|        | (1)<br>posresponse  | (2)<br>posresponse    |
|--------|---------------------|-----------------------|
| white  | 0.0242***<br>(4.30) | -0.000868<br>(-0.09)  |
| crime  | -0.0129*<br>(-2.36) | -0.0520***<br>(-4.23) |
| ged    | -0.00519<br>(-1.03) | 0.0105<br>(0.78)      |
| empgap | 0.00125<br>(0.27)   | 0.0103<br>(1.02)      |
| pre    | -0.0151<br>(-1.48)  |                       |
| N      | 14637               | 2918                  |

t statistics in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

(ii) Table (h) & (i)

The two tables are re-run and attached below, with the t-stat for each estimate shown.

Under observations with the 5% significance level, we can see all corresponding estimates remain consistent and share similar statistical significance conclusions across variables *callback* and *interview*. The results are quite similar, each of these analyses shows that racial discrimination increased substantially when these companies removed the box to comply with BTB.

```
. esttab mod21 mod22 mod23 mod25, drop(_cons *.center ged empgap)
```

|          | (1)<br>interview     | (2)<br>interview    | (3)<br>interview     | (4)<br>interview    |
|----------|----------------------|---------------------|----------------------|---------------------|
| boxwhite | -0.0262**<br>(-2.60) | -0.0409*<br>(-2.51) | -0.0360**<br>(-2.72) |                     |
| white    | 0.0169*<br>(2.41)    | 0.0399*<br>(2.20)   | 0.0323*<br>(2.24)    | 0.0158*<br>(2.33)   |
| box      | 0.0226<br>(1.46)     | 0.0108<br>(1.19)    | 0.00563<br>(0.68)    |                     |
| prewhite |                      |                     |                      | -0.00237<br>(-0.27) |
| pre      |                      |                     |                      | -0.00580<br>(-0.37) |
| N        | 7245                 | 3712                | 4794                 | 7476                |

t statistics in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

```
. esttab mod1 mod2 mod3 mod5, drop(_cons *.center ged empgap)
```

|          | (1)<br>posresponse  | (2)<br>posresponse  | (3)<br>posresponse  | (4)<br>posresponse |
|----------|---------------------|---------------------|---------------------|--------------------|
| boxwhite | -0.0302*<br>(-1.98) | -0.0364*<br>(-2.62) | -0.0327*<br>(-2.37) |                    |
| white    | 0.0315**<br>(2.63)  | 0.0442***<br>(3.51) | 0.0401**<br>(3.42)  | 0.0221*<br>(2.51)  |
| box      | 0.0150<br>(0.61)    | 0.00317<br>(0.21)   | -0.00175<br>(-0.13) |                    |
| prewhite |                     |                     |                     | 0.00203<br>(0.15)  |
| pre      |                     |                     |                     | -0.0159<br>(-0.91) |
| N        | 7245                | 3712                | 4794                | 7476               |

t statistics in parentheses  
\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Overall, to sum up, the paper's conclusions are similar for interview requests.



(k) After conducting temporal diff-in-diff analyses across column (2) to (4) for the same employers, the author commented that even if “the temporal diff-in-diff analysis is more causally rigorous than the cross-sectional comparison”, still “the estimates could be potentially confounded by trends unrelated to BTB”. And the author even mentioned an example about motivation to racial disparities in employment that might affect disparity trends in other ways. The analysis in column (5) was intended to check this concern, specifically column (5) was similar with column (2), but it only includes the sample of employers whose applications were unchanged/unaffected before and after BTB. Some estimate of interest is for the variable  $white \times pre$  and  $white$ . From column (5), the estimates are statistically insignificant, meaning time is insignificant in the relationship between the callback rates and race; also, there was essentially no change in the black-white gap between the pre-BTB and post-BTB periods.

However, if the estimate is negative and statistically significant, it would mean time has played an important role in the relationship between callback rate and race. That is, when there are employers unaffected after the BTB, they are more likely to call back to white applicant before the BTB than after BTB. Recall that the race gap was 3.6 percentage points smaller than after it was removed. Now, a negative and statistically significant result will confound our estimation, since the increase in race gap can be due to time, not BTB program itself. Hence, this assumption is problematic and will impede us from concluding causal relationships.