A164ps4

Matt Krasnow

April 2025

```
# Load data
pathways <- read_dta("Pathways To Education.dta")
```

Part 1: Pathways to Education

Question 1

```
# Regent Park comparison between cohorts
regent_data <- pathways %>% filter(regent_park == 1)
q1_reg <- lm(graduated5 ~ cohort >= 2001, data = regent_data)
summary(q1_reg)
```

```
##
## Call:
## lm(formula = graduated5 ~ cohort >= 2001, data = regent_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5929 -0.5929  0.4071  0.4071  0.6182
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.38182    0.03822   9.990  < 2e-16 ***
## cohort >= 2001TRUE   0.21108    0.04130   5.111 3.75e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4909 on 1148 degrees of freedom
## Multiple R-squared:  0.02225,    Adjusted R-squared:  0.0214
## F-statistic: 26.12 on 1 and 1148 DF,  p-value: 3.748e-07
```

1a. Interpretation: Students in Regent Park who entered 9th grade after the program was implemented (2001-2007) had graduation rates that were 21.1 percentage points higher than students who entered 9th grade before the program (2000). This difference is statistically significant ($p < 0.001$).

1b. Causal impact concerns: This difference might not measure the causal impact of the program because it doesn't account for other factors or general trends in graduation rates that might have affected all schools/housing projects during this time period, not just Regent Park.

Question 2

```
# DiD analysis excluding 2007 cohort
did_data <- pathways %>% filter(cohort <= 2006)
```

```r
did_data <- did_data %>% mutate(post = ifelse(cohort >= 2001, 1, 0))
q2_reg <- feols(graduated5 ~ regent_park*post,
                data = did_data,
                cluster = "project")
summary(q2_reg)
```

```
## OLS estimation, Dep. Var.: graduated5
## Observations: 5,847
## Standard-errors: Clustered (project)
##                   Estimate Std. Error  t value   Pr(>|t|)
## (Intercept)       0.441261   0.020289 21.74902  < 2.2e-16 ***
## regent_park      -0.059443   0.020289 -2.92983 4.5757e-03 **
## post              0.062712   0.021630  2.89931 4.9919e-03 **
## regent_park:post  0.142714   0.021630  6.59792 6.6985e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.498097   Adj. R2: 0.006992
```

2a. Interpretation: The difference-in-differences estimate of the program effect is 14.3 percentage points, indicating that the Pathways to Education program increased 5-year graduation rates in Regent Park by 14.3 percentage points relative to other housing projects. This effect is statistically significant ($p < 0.001$).

2b. Comparison with Question 1: This estimate is smaller than the estimate in Question 1 (14.3 vs 21.1 percentage points) because the DiD approach accounts for general trends in graduation rates affecting all housing projects during this time period, not just trends specific to Regent Park.

Question 3

```r
# TWFE regression
pathways <- pathways %>%
  mutate(treatment = case_when(
    regent_park == 1 & cohort >= 2001 ~ 1,
    (lawrence_heights == 1 | thistletown == 1) & cohort == 2007 ~ 1,
    TRUE ~ 0
  ))

q3_reg <- feols(graduated5 ~ treatment | cohort + project,
                data = pathways,
                cluster = "project")
summary(q3_reg)
```

```
## OLS estimation, Dep. Var.: graduated5
## Observations: 6,770
## Fixed-effects: cohort: 8,  project: 71
## Standard-errors: Clustered (project)
##           Estimate Std. Error t value   Pr(>|t|)
## treatment  0.12203   0.022283 5.47628 6.4006e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.490676     Adj. R2: 0.024455
##                  Within R2: 0.001698
```

3. Interpretation: The two-way fixed effects estimate indicates that the Pathways to Education program increased graduation rates by 12.2 percentage points across all treated locations. This effect is statistically significant ($p < 0.001$).

Question 4

```
# Extract cohort fixed effects
cohort_fe <- fixef(q3_reg)$cohort
cohort_fe
```

```
##      2000      2001      2002      2003      2004      2005      2006      2007
## 0.5430864 0.5474876 0.5747816 0.6312275 0.6392103 0.6156224 0.6383930 0.6792601
```

4. Graduation Rate Trends: Graduation rates are generally increasing over time, as evidenced by the increasing values of the cohort fixed effects from 2000 (0.543) to 2007 (0.679), with only a small dip in 2005.

Question 5

5. Variable controlled by cohort fixed effects: Economic conditions in Toronto in the year students entered 9th grade. This variable is controlled for by cohort fixed effects because it affects all students who entered 9th grade in the same year, regardless of which housing project they lived in. The assumption required is that these economic conditions had similar effects on students across all housing projects.

Question 6

6. Variable controlled by project fixed effects: Quality of the local elementary schools that feed into high schools from each housing project. This is controlled for by project fixed effects because it's a characteristic of the housing project that remains relatively constant over the study period. The assumption required is that the quality of these feeder schools did not change significantly during the years of the study.

Question 7

7a. Variable not controlled by either fixed effects: Changes in the quality or resources of individual high schools over time that specifically affect students from certain housing projects. This isn't controlled for because it varies both across projects (if students from different projects attend different high schools) and over time.

7b. Potential bias: This omitted variable could bias our treatment effect estimate if high schools serving Regent Park students improved their quality or resources around the same time the Pathways program was implemented. If this occurred, part of what we attribute to the Pathways program might actually be due to improvements in school quality. The bias would likely be positive (upward), making the program appear more effective than it actually was. However, if the program caused high schools to allocate more resources to Regent Park students, this would be part of the treatment effect we want to capture.