ARE 261 - Joe's Half
Problem Set 2

# Problem 1

Figure 1 plots total daily NOx emissions in the NBP-participating states.

# Problem 2

Note: According to Figure 1 in the paper, the NBP participating states include Alabama, Connecticut, Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, Massachusetts, Michigan, Missouri, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Virginia, and West Virginia. However, since Missouri only entered the market in 2007, I exclude it from the list of NBP states for the proceeding analysis since we are only utilizing 2002 and 2005 data.

### Part a

The econometric equation is

$$Y_d = \beta_0 + \beta_1 \text{Summer}_d + \beta_2 \tilde{X}_d + \beta_3 \tilde{X}_d^2 + \epsilon_t \tag{1}$$

where $Y_d$ is the total daily NOx emissions (in tons) on day-of-year $d$ in 2005, $\text{Summer}_d$ is an indicator equal to one for days between May 1st and September 30th (the ozone season), $\tilde{X}_d = \gamma_d - c$, where $\gamma_d$ is the day of the year and $c$ is the cutoff date - either May 1st or September 30th. The coefficient of interest, $\beta_1$, represents the estimated effect of the NOx Budget Trading Program on NOx emissions at the cutoff date.

### Part b

Columns 1–2 of Table 1 report the estimated effect of the NOx Budget Trading Program on NOx emissions using the polynomial RD. Using the May 1st (September 30th) cutoff, I find that the NBP led to an average decline in total NOx emissions of 2069.498 (2528.294) tons in 2005 from states participating in the NBP.[1] These estimates are statistically significant at the 1% level.

### Part c

Because we are summing total NOx emissions to the daily level in 2005, there is only variation across days within each regression's 2-month sample window. Thus, OLS is computing a difference in means at the cutoff date after flexibly controlling for quadratic trends across the sample period. This difference in means is unweighted and includes all observations in the 30 day window on either side of the cutoff. That is, we are using a rectangular kernel.

1. I interpreted "1 month on each side of the discontinuity" as ± 30 days (inclusive) on either side of the cutoff date.

This seems reasonable since we are summing emissions across states and only using one year of data. Should we desire, we could explore alternate kernel types (e.g,. triangular) to test the robustness of our specification.

The key identifying assumption in this design is continuity, which means that the expected potential outcomes are continuous at the cutoff. This assumption will fail if one or more of the determinants of NOx aside from the NBP (e.g., weather flucuations or economic activity) jumps discontinuously at the cutoff. Our RD estimates will be biased if the continuity assumption is not satisfied.

Lastly, I use heteroskedasticity robust standard errors to obtain unbiased standard errors of the OLS coefficients under heteroscedasticity.

# Problem 3

### Part a

$$Y_d = \beta_0 + \beta_1 \text{Summer}_d + \beta_2 \tilde{X}_d + \beta_3 \tilde{X}_d^2 + \beta_4 \tilde{X}_d \times \text{Summer}_d + \beta_5 \tilde{X}_d^2 \times \text{Summer}_d + \epsilon_d, \quad (2)$$

where $\text{Summer}_d$ and $\tilde{X}_d$ are defined as before.

### Part b

Columns 3–4 of Table 1 report the estimated effect of the NOx Budget Trading Program on NOx emissions using the spline RD. Using the May 1st (September 30th) cutoff, I find that the NBP led to a decline in total NOx emissions of roughly 1835.393 (2701.130) tons in 2005 from states participating in the NBP. These estimates are statistically significant at the 1% level.

# Problem 4

### Part a

The econometric equation is

$$Y_d = \beta_0 + \beta_1 \text{Summer}_d + \epsilon_d, \quad (3)$$

where $Y_d$ is total daily NOx emissions (in tons) on day-of-year $d$ in 2005, $\text{Summer}_d$ is defined as before, and $\epsilon_d$ is an error term. The coefficient of interest, $\beta_1$, represents the difference in 2005 average total daily NOx emissions in the summer versus the winter from NBP participating states. That is, $\hat{\beta}_1 = (\bar{Y}_{\text{Summer}} - \bar{Y}_{\text{Winter}})$.

### Part b

Using this cross-sectional estimator in 2005, I estimate that the NBP reduced total daily summer NOx emissions in 2005 by an average of 3097.434 tons (Column 1, Table 2). This estimate is statistically significant at the 1% level.

# Problem 5

## Part a

The econometric equation is

$$Y_{dt} = \beta_0 + \beta_1 \text{Summer}_d + \beta_2 \text{Post}_t + \beta_3 \text{Summer}_d \times \text{Post}_t + \epsilon_{dt}, \tag{4}$$

where $Y_{dt}$ is total daily NOx emissions (in tons) on day of year $d$ and year $t$, $\text{Summer}_d$ is defined as before, $\text{Post}_t$ is a binary indicator equal to one for the year 2005 and zero for the year 2002, and $\epsilon_{dt}$ is an error term. The coefficient of interest, $\beta_3$, represents the difference in average total daily NOx emissions in the summer vs winter and before vs after the NBP went into effect in NBP participating states. That is,

$$\hat{\beta}_3 = (\bar{Y}_{\text{Summer,Post}} - \bar{Y}_{\text{Winter,Post}}) - (\bar{Y}_{\text{Summer,Pre}} - \bar{Y}_{\text{Winter,Pre}}).$$

## Part b

Using the pre/post DiD estimator, I estimate that the NBP reduced total daily summer NOx emissions in 2005 by an average of 2888.387 tons (Column 2, Table 2). This estimate is statistically significant at the 1% level.

# Problem 6

## Part a

The econometric equation is

$$Y_{id} = \beta_0 + \beta_1 \text{Summer}_d + \beta_2 \text{East}_i + \beta_3 \text{Summer}_d \times \text{East}_i + \epsilon_{id}, \tag{5}$$

where $Y_{id}$ is total daily NOx emissions (in tons) in region $i$ (East vs West) on day-of-year $d$, $\text{Summer}_d$ is defined as before, $\text{East}_i$ is a binary indicator equalling one for Eastern states participating in the NBP and zero for Western states, and $\epsilon_{id}$ is an error term. The coefficient of interest, $\beta_3$, represents the difference in 2005 average total daily NOx emissions in the summer vs winter and in the East vs West.[2] That is,

$$\hat{\beta}_3 = (\bar{Y}_{\text{Summer,East}} - \bar{Y}_{\text{Winter,East}}) - (\bar{Y}_{\text{Summer,West}} - \bar{Y}_{\text{Winter,West}}).$$

## Part b

Using the East/West DiD, I estimate that the NBP reduced total daily summer NOx emissions in 2005 by an average of 3496.941 tons (Column 3, Table 2). This estimate is statistically significant at the 1% level.

2. Following the main analysis of the paper, I exclude Wisconsin, Iowa, Missouri, Georgia, Mississippi, Maine, New Hampshire, and Vermont from the regression sample.

# Problem 7

### Part a

The econometric equation is

$$\begin{aligned} Y_{idt} = \beta_0 &+ \beta_1 \text{Summer}_d + \beta_2 \text{East}_i + \beta_3 \text{Post}_t + \beta_4 \text{Summer}_d \times \text{East}_i + \beta_5 \text{Summer}_d \times \text{Post}_t \\ &+ \beta_6 \text{East}_i \times \text{Post}_t + \beta_7 \text{Summer}_d \times \text{East}_i \times \text{Post}_t + \epsilon_{idt}, \end{aligned} \tag{6}$$

where $Y_{idt}$ is total NOx emissions (in tons) in region $i$ (East vs West), day of year $d$, and year $t$, Summer$_d$, East$_i$, and Post are defined as before, and $\epsilon_{idt}$ is an error term. The coefficient of interest, $\beta_7$, represents the difference in average total daily NOx emissions in the East vs West, summer vs winter, and before vs after the NBP went into effect. That is,

$$\begin{aligned} \hat{\beta}_7 = &[(\bar{Y}_{\text{Summer,East,Post}} - \bar{Y}_{\text{Summer,East,Pre}}) - (\bar{Y}_{\text{Winter,East,Post}} - \bar{Y}_{\text{Winter,East,Pre}}] \\ &- [(\bar{Y}_{\text{Summer,West,Post}} - \bar{Y}_{\text{Summer,West,Pre}}) - (\bar{Y}_{\text{Winter,West,Post}} - \bar{Y}_{\text{Winter,West,Pre}})]. \end{aligned}$$

### Part b

Using the triple difference estimator, I estimate that the NBP reduced total daily summer NOx emissions by an average of 2911.723 tons (Column 3, Table 2). This estimate is statistically significant at the 1% level.

### Part c

I choose to not weight the regressions. I do this because we are interested in the effect of the NBP on average total daily emissions across units affected by the NBP. Although I do not cluster the standard errors, we could consider doing so if we are worried about serial or spatial autocorrelation. For instance, if we were running a state-level season regression rather than an east vs west daily regression, I could consider clustering at the state-season level.

# Problem 8

For the triple difference estimator in Problem 7, the NBP should affect Eastern states only: $D_{idt} = 1[i = \text{East}] \times 1[d \in \text{Summer}] \times 1[t = 2005]$. This assumption might fail if there are spillover effects onto Western states (e.g., wind from the East to the West carrying fewer NOx emissions than usual), although the authors select the control states with this concern in mind. In the DDD estimator, we also allow for non-parallel trends between the East vs West, but require this non-parallelness to be the same during the summer vs winter. That is,

$$\mathbb{E}[\Delta Y_{\text{East, Summer}}(0) - \Delta Y_{\text{West, Summer}}(0)] = \mathbb{E}[\Delta Y_{\text{East, Winter}}(0) - \Delta Y_{\text{West, Winter}}(0)]$$

This assumption might fail if there are unobserved seasonal shocks to NBP states after the NBP went into effect (e.g., other regulatory changes that occur during the summer months in NBP states in 2005).

I believe the DDD estimator is the most reasonable approach to measuring the effects of this program. The polynomial RD in Problem 2 only uses one year of (post-period) data and, thus, cannot rule out the possibility that NOx emissions in NBP states typically fall during the summer months (specifically, around the cutoff dates) even before the program went into effect, in which case there would be no treatment effect. This specification also imposes a common quadratic time trend on either side of the cutoff dates, which may be a stronger assumption than we are comfortable making. Although the spline RD in Problem 3 allows for different quadratic trends on either side of the cutoff dates, it suffers from the first issue that afflicts the polynomial RD.

Both RD specifications are likely better identified than the cross-sectional specification in Problem 4 since they are comparing outcomes around 2-month windows of the program start/stop rather than across summer vs winter. The months on either side of the cutoff dates are likely more comparable than summer vs winter. However, the RD specifications only return the local average treatment effect (LATE) during these 2-month windows, whereas we are presumably interested in the average treatment effect on the treated (ATT) for the entire summer period when the NBP operates.

Lastly, the parallel trends assumption in the DDD estimator is weaker than those in the DiD estimators, making it more likely to be satisfied in our setting. Specifically, the DiD estimator in Problem 5 requires parallel summer vs winter trends among NBP states before vs after the NBP began operating in 2005:

$$\mathbb{E}[Y_{\text{Summer, Post}}(0) - Y_{\text{Winter, Post}}(0)] = \mathbb{E}[Y_{\text{Summer, Pre}}(0) - Y_{\text{Winter, Pre}}(0)],$$

and the DiD estimator in Problem 6 requires parallel summer vs winter trends in 2005 in the East vs West:

$$\mathbb{E}[Y_{\text{Summer, East}}(0) - Y_{\text{Winter, East}}(0)] = \mathbb{E}[Y_{\text{Summer, West}}(0) - Y_{\text{Winter, West}}(0)].$$

# Problem 9

My interpretation of this question is that, instead of running the above regressions to estimate the effects of the NBP program, we are proposing comparing pre-period measured emissions with post-period emissions as prescribed by the caps. If the caps were perfectly binding and we assumed that firms would want to emit as much pollution as they could, then this approach would tell us how much emissions fell during the summer when the NBP is in effect. However, since the NBP is not in effect during the winter, using the caps along will not tell us how emissions changed during that period of time. Moreover, since firms may emit beyond or below their allowances, the prescribed emissions caps will almost certainly not match the actual level of emissions once the NBP went into effect.

# Problem 10

The consumer's problem is

$$\max_{X,f,a} u(X, f, s(c, a)) \text{ s.t. } I + p_w(T - f - s(c, a)) \geq X + p_a a,$$

where $X$ is the numeraire good, $f$ is hours of leisure, and $s = s(c, a)$ is health as captured by the number of sick days, which depends on the ambient pollution concentration $c$ and defensive behavior $a$. The budget constraint is comprised of non-labor income, $I$, and labor income, $p_w(T - f - s(c, a))$. We can therefore write the Lagrangian as

$$\mathcal{L} = u(X, f, s(c, a)) + \lambda(I + p_w(T - f - s(c, a)) - X + p_a a)$$

with first-order conditions

$$\frac{\partial u}{\partial X} = \lambda, \tag{7}$$

$$\frac{\partial u}{\partial f} = \lambda p_w \tag{8}$$

$$\frac{\partial u}{\partial s}\frac{\partial s}{\partial a} = \lambda\left[\frac{\partial s}{\partial a}p_w + p_a\right] \tag{9}$$

Recall that the marginal cost of the numeraire good, $X$ , is 1; the marginal cost of leisure, $f$, is $p_w$ (i.e., the hourly wage you give up by not working); and the marginal cost of defensive actions, $a$, is the direct cost, $p_a$, net of the saved costs from taking fewer sick days (i.e., the additional wages that can be earned from taking fewer sick days due to improved health) since $\partial s/\partial a < 0$. Moreover, recall that $\lambda$ is the "shadow price," or the utility gain at the optimum of relaxing the budget constraint by one dollar. Therefore, each first-order condition says that the marginal utility of the good in question equals the shadow price times the marginal cost. In other words, the marginal utility of consuming one additional unit equals the marginal disutility of purchasing one additional unit.

# Problem 11

Given that the consumer chooses $a$ to maximize their utility, we can write their health production function as $s = s(c, a^*)$. Totally differentiating this with respect to $s$ yields

$$\frac{ds}{dc} = \frac{\partial s}{\partial c} + \frac{\partial s}{\partial a}\frac{\partial a^*}{\partial c}$$

Rearranging terms, we get

$$\frac{\partial s}{\partial c} = \frac{ds}{dc} - \frac{\partial s}{\partial a}\frac{\partial a^*}{\partial c}. \tag{10}$$

# Problem 12

$\partial s/\partial c$ represents the partial effect of ambient pollution on health as measured by sick days, holding all else equal. $ds/dc$ represents the total effect of ambient pollution concentration on health, which is the direct effect of pollution on health net of the mitigating effects of defensive behaviors.

$\partial s/\partial c$ is difficult to estimate both in an experimental and non-experimental setting. In an experimental setting, we would need to randomly expose some people to higher ambient pollution concentrations without allowing them to take defensive actions, which is highly unethical. In the non-experimental case, we would need data on all defensive investments according to the second RHS term in equation (10), which is typically more difficult to obtain than pollution or health data. $ds/dc$ is comparatively easier to estimate because observed health outcomes are already net of defensive behaviors people engage in.

# Problem 13

First, we can rearrange equation (10) as

$$\frac{\partial s}{\partial a} = \frac{1}{\partial a^*/\partial c}\left[\frac{ds}{dc} - \frac{\partial s}{\partial c}\right]$$

Substituting this expression for $\partial s/\partial a$ into equation (9), we get

$$\frac{\partial u}{\partial s}\left(\frac{1}{\partial a^*/\partial c}\left[\frac{ds}{dc} - \frac{\partial s}{\partial c}\right]\right) = \lambda p_w\left(\frac{1}{\partial a^*/\partial c}\left[\frac{ds}{dc} - \frac{\partial s}{\partial c}\right]\right) + \lambda p_a$$

Simplifying and rearranging yields

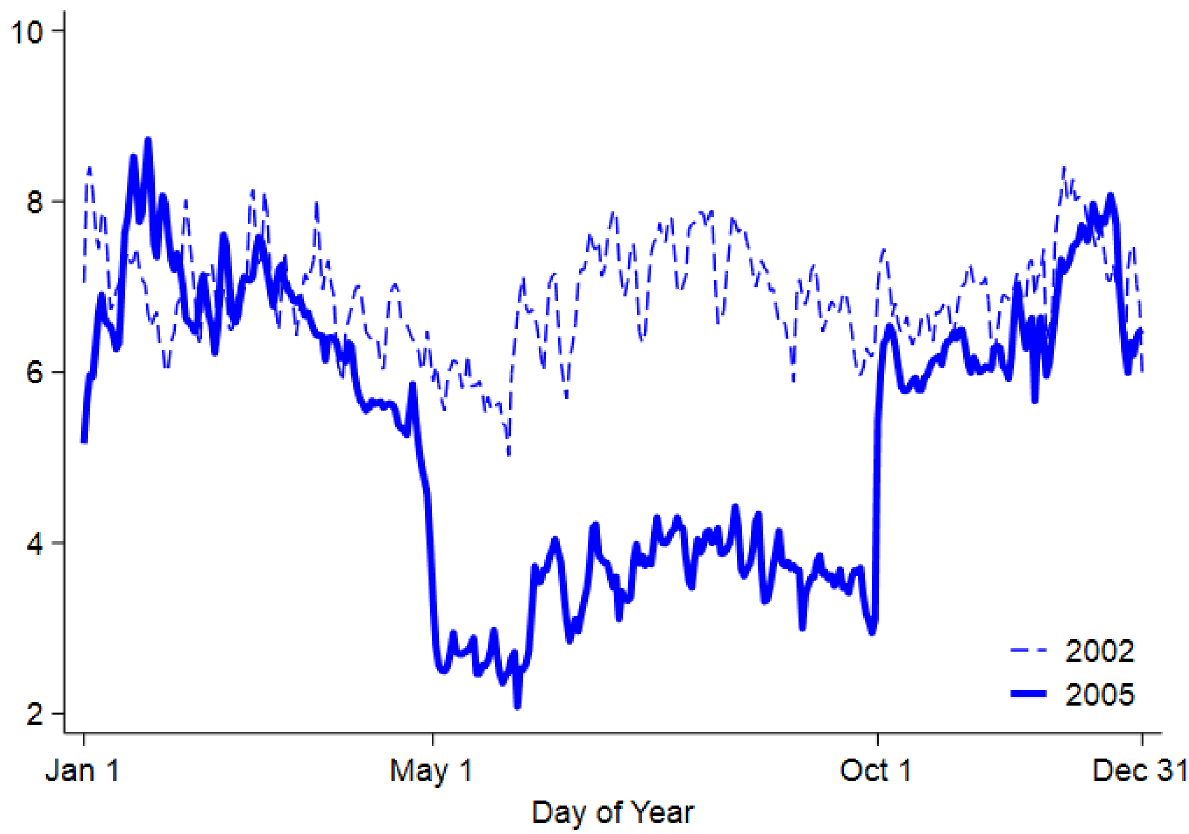
$$\begin{aligned}
&\frac{\partial u}{\partial s}\left[\frac{ds}{dc} - \frac{\partial s}{\partial c}\right] = \lambda p_w\left[\frac{ds}{dc} - \frac{\partial s}{\partial c}\right] + \lambda p_a\frac{\partial a^*}{\partial c}\\
\Leftrightarrow\ &\frac{\partial u}{\partial s}\frac{ds}{dc} - \frac{\partial u}{\partial s}\frac{\partial s}{\partial c} = \lambda p_w\frac{ds}{dc} - \lambda p_w\frac{\partial s}{\partial c} + \lambda p_a\frac{\partial a^*}{\partial c}\\
\Leftrightarrow\ &\lambda p_w\frac{\partial s}{\partial c} - \frac{\partial u}{\partial s}\partial s\partial c = \lambda p_w\frac{ds}{dc} + \lambda p_a\frac{\partial a^*}{\partial c} - \frac{\partial u}{\partial s}\frac{ds}{dc}\\
\Leftrightarrow\ &p_w\frac{\partial s}{\partial c} - \frac{1}{\lambda}\frac{\partial u}{\partial s}\partial s\partial c = \left(p_w\frac{ds}{dc}\right) + \left(p_a\frac{\partial a^*}{\partial c}\right) - \left(\frac{\partial u/\partial s}{\lambda}\frac{ds}{dc}\right)\\
\Leftrightarrow\ &w_c = \left(p_w\frac{ds}{dc}\right) + \left(p_a\frac{\partial a^*}{\partial c}\right) - \left(\frac{\partial u/\partial s}{\lambda}\frac{ds}{dc}\right).
\end{aligned} \tag{11}$$

# Problem 14

In the model, $p_a$ represents the cost of all defensive behaviors taken against air pollution. It is worth noting that medications alone - which is what the paper uses to measure defensive investments - do not capture the full spectrum of potential defensive investments (e.g., purchasing air filters or limiting time spent outdoors).

# Figures

Figure 1: Total Daily NOX Emissions in the NBP-Participating States




*Notes:* Figure 1 shows average total daily NOx emissions in the NBP participating states in 2002 and 2005. These estimates are obtained from an OLS regression of NOx emissions on 6 day-of-week indicators and a constant. The values in the graph equal the constant plus the regression residuals, so that the graph depicts fitted values for the reference category (Wednesday). Total daily NOx emissions on Y-axis are measured in thousands of tons. The sample includes emissions from all the Acid Rain Units. NBP participating states include: Alabama, Connecticut, Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, Massachusetts, Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Virginia, and West Virginia. Note that I do not include Missouri since the NBP went into effect in 2007 there.

# Tables

Table 1: Polynomial and Spline RD estimates for the effect of the NBP on NOx emissions

| | Polynomial RD | | Spline RD | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | NOx | NOx | NOx | NOx |
| 1(NBP Operating) | -2069.498*** | -2528.294*** | -1835.393*** | -2701.130*** |
| | (217.871) | (249.646) | (233.808) | (315.385) |
| Cutoff date | May 1 | Sep. 30 | May 1 | Sep. 30 |

*Notes*: The table dislpays estimates for the the effect of the NOx Budget Trading Program on average total daily NOx emissions. Columns 1–2 report the results using the polynomial regresion discontinuity, and Columns 3–4 report the results using the spline regression discontinuity. Columns 1 and 3 include the 30 days before and after May 1st in the sample, and Columns 2 and 4 include the 30 days before and after September 30th. Robust errors are in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 2: Cross-sectional and DiD estimates of the effect of the NBP on NOx emissions

| (1) | (2) | (3) | (4) |
|---|---|---|---|
| -3097.434*** | -2888.387*** | -3496.941*** | -2911.723*** |
| (73.936) | (107.501) | (80.955) | (117.365) |

*Notes:* The table dislpays estimates for the the effect of the NOx Budget Trading Program on average total daily NOx emissions. Columns 1 reports the estimated coefficient of interest from the cross-sectional specification in Question 4. Column 2 the estimated coefficient of interest from the Pre vs Post DiD in Question 5. Column 3 reports the estimated coefficient of interest from the East vs West DiD in Question 6. Finally, Column 4 reports the estimated coefficient of interest from the triple-difference specification in Question 7. Robust errors are in parentheses. * p<0.05, ** p<0.01, *** p<0.001

# Code

### Setup

```
********************************************************************************
* ARE 261 Pset 1 - Joe's Half
* Last updated: October 22, 2023
* Purpose: initialize file paths and settings
********************************************************************************

clear all
set more off
set scheme scheme_fb2

global dirpath_home "\\tsclient\Documents\github\ARE-261-Problem-Sets\Pset3"
global dirpath_server "C:\Users\gschlauch\Documents\ARE261\pset3"
global dirpath_code "$dirpath_home\code"
global dirpath_data "$dirpath_server\data"
global dirpath_output "$dirpath_home\output"
```

### Clean data

```
********************************************************************************
* Purpose: clean the raw data
********************************************************************************

* Initialize settings and filepaths
do "\\tsclient\Documents\github\ARE-261-Problem-Sets\Pset3\scripts\setup.do"

* Load the raw data
import delimited "$dirpath_data\raw\raw_emissions_data.csv", clear

* Keep key variables
keep statecode facilityname facilityid unitid date noxmass
rename statecode state
rename facilityname facility_name
rename facilityid facility_id
rename unitid unit_id
rename noxmass nox_mass

* Check that the data are uniquely identified as expected
gunique state facility_id unit_id date
assert r(unique) == _N

* Create stata date variables
gen date_stata = date(date, "YMD")
format date_stata %td
gen year = year(date_stata)
drop date

* Conver the NOx mass variable to numeric
```

```
replace nox_mass = "" if nox_mass == "NA"
destring nox_mass, replace

* Create NBP binary indicator = 1 for NBP states and 0 otherwise. Note that I
* exclude Missouri from the list because the NBP did not begin operating there
* until 2007 and the problem set is focused on 2002 vs 2005.
gen nbp = 0
foreach stabv in AL CT DE DC IL IN KY MD MA MI NJ NY NC OH PA RI SC TN VA WV {
    qui replace nbp = 1 if state == "‘stabv’"
}

* Create indicator for eastern states, excluding states that were excluded in
* the paper (including Missouri, since the NBP did not begin operating there
* until 2007). I use this indicator later when I run east vs west regressions
gen east = nbp
foreach stabv in WI IA MO GA MS ME NH VT AK HI {
    qui replace east = . if state == "‘stabv’"
}

* Output
compress *
save "$dirpath_data\clean\NOx_data_cleaned.dta", replace
```

### Analysis

```
****************************************************************************
* Purpose: analyze the cleaned data
****************************************************************************

* Initialize settings and filepaths
do "\\tsclient\Documents\github\ARE-261-Problem-Sets\Pset3\scripts\setup.do"

* Question 1 ******************************************************************

use "$dirpath_data\clean\NOx_data_cleaned.dta", clear

* Keep NBP participating states
keep if nbp == 1

* Collapse the emissions data by day-year
gcollapse (sum) nox_mass, by(year date_stata)

* Convert to thousands of tons
replace nox_mass = nox_mass / 1000

* Create day of week indicators
gen dow = dow(date_stata)
tab dow, gen(dow)

* Regress NOx emissions on 6 day-of-week indicators and a constant
reghdfe nox_mass dow1-dow3 dow5-dow7 dow4, noabs residuals(resid_nox_emit)
gen fit_nox_emit = resid_nox_emit + _b[_cons]
```

```
* Plot Figure 1
gen doy = doy(date_stata)
gsort year doy
twoway ///
    (line fit_nox_emit doy if year == 2002, ///
        lcolor(blue) lpattern(dash)) ///
    (line fit_nox_emit doy if year == 2005, ///
        lcolor(blue) lpattern(solid) lwidth(thick)), ///
    xtitle("Day of Year") ///
    legend(order(1 "2002" 2 "2005") pos(5) ring(0) rows(2)) ///
    xlab(1 "Jan 1" 121 "May 1" 274 "Oct 1" 365 "Dec 31") ///
    ytitle("")
graph export "$dirpath_output\Figures\Fig1.png", replace


* Questions 2 and 3 **********************************************************

clear all
use "$dirpath_data\clean\NOx_data_cleaned.dta"

* Keep NBP participating states
keep if nbp == 1

* Restrict to the year 2005
keep if year == 2005

* Get the total emissions by date
gcollapse (sum) nox_mass, by(date_stata)
gen month = month(date_stata)

* Create treatment indicator = 1 during ozone season
gen summer = inlist(month, 5, 6, 7, 8, 9)

* Create RD window indicators
gen rdwindow_1 = (inrange(date_stata, td(01may2005) - 30, td(01may2005) + 30))
gen rdwindow_2 = (inrange(date_stata, td(30sep2005) - 30, td(30sep2005) + 30))

* Create running variables cenetered at the cutoff values
gen runvar_minus_c1 = date_stata - td(01may2005)
gen runvar_minus_c1_sq = runvar_minus_c1^2

gen runvar_minus_c2 = date_stata - td(30sep2005)
gen runvar_minus_c2_sq = runvar_minus_c2^2

* Run the regressions for question 2
forvalues i = 1/2 {
    eststo: reghdfe nox_mass summer ///
        runvar_minus_c`i' runvar_minus_c`i'_sq ///
        if rdwindow_`i' == 1, noabs vce(robust)
    if `i' == 1 {
        estadd local cutoff "May 1"
    }
    else {
```

```
        estadd local cutoff "Sep. 30"
    }
}

* Run the regressions for question 3
forvalues i = 1/2 {
    eststo: reghdfe nox_mass summer ///
        runvar_minus_c‘i’ runvar_minus_c‘i’_sq ///
        c.runvar_minus_c‘i’#i.summer ///
        c.runvar_minus_c‘i’_sq#i.summer ///
        if rdwindow_‘i’ == 1, noabs
    if ‘i’ == 1 {
        estadd local cutoff "May 1"
    }
    else {
        estadd local cutoff "Sep. 30"
    }
}

* Create latex table containing the point estimates on summer
cd "$dirpath_output\Tables"
la var nox_mass "NOx"
la var summer "1(NBP Operating)"

local longnote "\emph{Notes}: The table dislpays estimates for the the effect of the NOx
    Budget Trading Program on average total daily NOx emissions. Columns 1--2 report the
    results using the polynomial regresion discontinuity, and Columns 3--4 report the
    results using the spline regression discontinuity. Columns 1 and 3 include the 30
    days before and after May 1st in the sample, and Columns 2 and 4 include the 30 days
    before and after September 30th. Robust errors are in parentheses. * p<0.05, **
    p<0.01, *** p<0.001"

esttab using "Table_RD_estimates.tex", replace ///
    title("Polynomial and Spline RD estimates for the effect of the NBP on NOx emissions
        \label{tab1}") ///
    label b(3) se(3) keep(summer) ///
    mgroups( ///
        "Polynomial RD" "Spline RD", pattern(1 0 1 0) span ///
        prefix(\multicolumn{@span}{c}{) suffix(}) ///
        erepeat(\cmidrule(lr){@span}) ///
        ) ///
    substitute(\_ _ {l} {p{0.8\linewidth}}) wrap ///
    stats(cutoff, labels("Cutoff date") fmt(0)) booktabs ///
    nonotes addnotes("‘longnote’")
eststo clear


* Question 4 *******************************************************************

use "$dirpath_data\clean\NOx_data_cleaned.dta", clear
keep if year == 2005
keep if nbp == 1
gen month = month(date_stata)
gen summer = inlist(month, 5, 6, 7, 8, 9)
```

```
gcollapse (sum) nox_mass, by(date_stata summer)

reghdfe nox_mass summer, noabs vce(robust)

local var1 summer
local b1 = _b[`var1']
local se1 = _se[`var1']

* Question 5 ***************************************************************

use "$dirpath_data\clean\NOx_data_cleaned.dta", clear
keep if east == 1
gen month = month(date_stata)
gen summer = inlist(month, 5, 6, 7, 8, 9)
gcollapse (sum) nox_mass, by(date_stata year summer)
gen post = (year == 2005)
gen summerXpost = summer * post

reghdfe nox_mass summer post summerXpost, noabs vce(robust)

local var2 summerXpost
local b2 = _b[`var2']
local se2 = _se[`var2']

* Question 6 ***************************************************************

use "$dirpath_data\clean\NOx_data_cleaned.dta", clear
keep if year == 2005
drop if missing(east)
gen month = month(date_stata)
gen summer = inlist(month, 5, 6, 7, 8, 9)
gcollapse (sum) nox_mass, by(date_stata east summer)
gen summerXeast = summer * east

reghdfe nox_mass summer east summerXeast, noabs vce(robust)

local var3 summerXeast
local b3 = _b[`var3']
local se3 = _se[`var3']

* Question 7 ***************************************************************

use "$dirpath_data\clean\NOx_data_cleaned.dta", clear
drop if missing(east)
gen month = month(date_stata)
gen summer = inlist(month, 5, 6, 7, 8, 9)
gcollapse (sum) nox_mass, by(date_stata year east summer)
gen post = (year == 2005)
gen summerXeast = summer*east
gen summerXpost = summer*post
gen eastXpost = east*post
gen summerXeastXpost = summer*east*post

reghdfe nox_mass summer post east summerXeast summerXpost ///
```

```
    eastXpost summerXeastXpost, noabs vce(robust)

local var4 summerXeastXpost
local b4 = _b[‘var4’]
local se4 = _se[‘var4’]

* Tabulate the estimates from Questions 4-7 **********************************

* Get the significance stars for the parameter estimate of interest in each
* regression
forvalues i = 1/4 {

    local t_stat = ‘b‘i’’ / ‘se‘i’’
    local pval = 2 * ttail(e(df_r), abs(‘t_stat’))
    if ‘pval’ < 0.01 {
        local stars‘i’ = "***"
    }
    else if ‘pval’ < 0.05 {
        local stars‘i’ = "**"
    }
    else if ‘pval’ < 0.1 {
        local stars‘i’ = "*"
    }
    else {
        local stars‘i’ = ""
    }

}

* Begin table
local own_file = 0
capture file close myfile
file open myfile using
    "$dirpath_output/tables/Table_cross-sectional_and_DiD_estimates.tex", write replace
if ‘own_file’ == 1 {
file write myfile "\documentclass[12pt]{article}" _n
file write myfile "\usepackage{amsmath}" _n
file write myfile "\usepackage{tabularx}" _n
file write myfile "\usepackage{booktabs}" _n
file write myfile "\begin{document}" _n
file write myfile "\pagenumbering{gobble}" _n
file write myfile _n
}
file write myfile "\begin{table}[ht]" _n
file write myfile "\caption{Cross-sectional and DiD estimates of the effect of the NBP
    on NOx emissions}" _n
file write myfile "\centering" _n
file write myfile "\normalsize" _n
file write myfile "\begin{tabular}{cccc}" _n
file write myfile "\toprule" _n
file write myfile "\centering" _n
file write myfile " (1) & (2) & (3) & (4) \\" _n
file write myfile "\midrule" _n
```

```
* Write the results
forvalues i = 1(1)4 {
    local b`i' = round(`b`i'', 0.001)
    file write myfile "`b`i''`stars`i''" _tab
    if `i' != 4 {
        file write myfile " &" _tab
    }
}
file write myfile "\\" _n

forvalues i = 1(1)4 {
    local se`i' = round(`se`i'', 0.001)
    file write myfile "(`se`i'')" _tab
    if `i' != 4 {
        file write myfile " &" _tab
    }
}
file write myfile "\\" _n

* End table
file write myfile "\bottomrule" _n
file write myfile "\end{tabular}" _n
file write myfile "\caption*{\footnotesize \emph{Notes:} The table dislpays estimates
    for the the effect of the NOx Budget Trading Program on average total daily NOx
    emissions. Columns 1 reports the estimated coefficient of interest from the
    cross-sectional specification in Question 4. Column 2 the estimated coefficient of
    interest from the Pre vs Post DiD in Question 5. Column 3 reports the estimated
    coefficient of interest from the East vs West DiD in Question 6. Finally, Column 4
    reports the estimated coefficient of interest from the triple-difference
    specification in Question 7. Robust errors are in parentheses. * p<0.05, ** p<0.01,
    *** p<0.001}" _n
file write myfile "\label{table:CS_DiD}" _n
file write myfile "\end{table}" _n
if `own_file' == 1 {
file write myfile "\end{document}" _n
}
file close myfile
```