

EEP/IAS 118 - Problem Set 5

Due Friday, December 3 at 11:59PM.

Submit materials as one combined pdf on Gradescope. All work can be completed in this notebook. Make sure to run (shift + enter) all your answer cells before submission to make sure all your output is displayed. After exporting your file to PDF, make sure that your output cells are not being cut off so we can read all of your code and results. If your output is getting cut off, try different ways of generating a PDF (File->Download as->PDF via HTML; go to print in your browser and save as PDF; whatever has worked for you or your peers in the past).

Exercise 1. The Effect of Minimum Wage on Employment - Difference-in-Differences

Background

In this exercise, we are going to look at a classic paper in the labor economics literature (note that we are only giving you a subset of the data, so your results are not going to match the results in the paper). This paper answers a very important (and often controversial) question in economic policy: does increasing the minimum wage increase unemployment (or conversely, reduce employment)? Proponents of minimum wage laws point to the benefits for individuals who remain employed in low wage jobs. Opponents of minimum wage laws argue that the increases in labor costs result in higher unemployment because employers hire less employees to offset the cost increases. Card and Krueger (1994) test this latter hypothesis using the minimum wage increase in New Jersey that went into effect in 1992. They surveyed fast food establishments in New Jersey and Pennsylvania both before and after the policy came into effect, collecting information on wages, employment and prices. Pre-policy change interviews are coded as 1992 and post-policy change interviews are coded as 1993 in your data for the sake of simplicity. This data is then used to obtain a difference-in-differences estimate of the effect of minimum wage laws on employment.

The dataset is saved as minwage_data.csv and contains the following variables:

| Variable Name | Description |
|---------------|--|
| store_id | Unique Store ID |
| year | Year |
| state | Dummy =1 if store is located in New Jersey, =0 for Pennsylvania |
| shore | Dummy =1 if store is located on New Jersey Shore, =0 otherwise |
| emply | Number of full-time employees in a store |
| emplyt | Number of part-time employees in a store |
| nmgrs | Number of managers in a store |
| wage_st | Starting wage in a store |
| penrce | Price of an entree |
| fte | Number of full-time equivalent employees in a store (emply + 0.5*emplyt + nmgrs) |

Question 1.1.

Load the data. Generate a summary table with two columns and two rows. There should be two columns: one for New Jersey (Treatment column) and one for Pennsylvania (Control column) and two rows: one for the pre-period (year 1992), one for the post-period (year 1993). Within each cell, compute the mean number of full-time equivalent employees (the variable fte).

Hint: Remember that you can subset data by writing, for example data[data\$var1==0 & data\$var2==7,]\$var3 to select values for var3 for observations in data that meet the given criteria for var1 and var2. The command cbind may be helpful for constructing your table. Remember that you can create a vector of values using c().

Hint: Consider loading necessary packages for the rest of the assignment here as well. It is good practice to load all necessary packages at the beginning of your code. Think about what packages we have needed previously in this class. You will also need the life package.

```
In [1]: install.packages('life')
library(tidyverse)
library(haven)
library(life)

mw <- read.csv("minwage_data.csv")

Installing package into 'C:/Users/pierrebliscaye/Library/R/x86_64/4.1/library'
(as 'lib' is unspecified)

also installing the dependency 'sandwich'
```

The downloaded binary packages are in
C:/var/folders/8z/9r58kd2d3hx5csf23tpt3d_80000gn/T/RtmpmkwXpT/downloaded_packages

```
--- Attaching packages --- tidyverse 1.3.1 ---

✔ ggplot2 3.3.5      ✔ dplyr 0.3.4
✔ tidbale 3.1.3      ✔ dplyr 1.0.7
✔ tidyr 1.1.3        ✔ stringr 1.4.0
✔ readr 2.0.1        ✔ forcats 0.5.1
```

```
--- Conflicts --- tidyverse_conflicts() ---
* dplyr::filter() masks stats::filter()
* dplyr::lag() masks stats::lag()

Loading required package: Matrix

Attaching package: 'Matrix'
```

The following objects are masked from 'package:tidyr':

expand, pack, unpack

```
In [2]: Year<-c("1992","1993")
PA<-c(mean(mw$state==0&mw$year==1992,$fte, na.rm=T),mean(mw$state==0&mw$year==1993,$fte, na.rm=
T))
NJ<-c(mean(mw$state==1&mw$year==1992,$fte, na.rm=T),mean(mw$state==1&mw$year==1993,$fte, na.rm=
T))

cbind(Year, PA, NJ)
```

A matrix: 2 x 3 of type chr

| | Year | PA | NJ |
|------|------------------|------------------|----|
| 1992 | 23.6727272727273 | 20.7566810311942 | |
| 1993 | 20.1741071428571 | 19.0855769230769 | |

Question 1.2.

State the difference-in-differences estimator for the change in full-time equivalent employees in terms of the following quantities $\bar{Y}_{NJ,pre}$, $\bar{Y}_{NJ,post}$, $\bar{Y}_{Penn,pre}$, $\bar{Y}_{Penn,post}$ where \bar{Y} refers to the mean of fte (writing a formula in R code is ok). Using the means reported in part 1, calculate a value for the estimator you just proposed.

```
In [3]: njpre <- NJ[1]
njpost <- NJ[2]
papre <- PA[1]
papost <- PA[2]

DD <- (njpost - njpre) - (papost - papre)
paste0("Our difference-in-differences estimator is ", DD)
```

'Our difference-in-differences estimator is 1.82751602175283'

Question 1.3.

Let's proceed with estimating the difference-in-differences estimator via a regression:

(a) Write an equation that will give you the difference-in-differences estimator for the impact of the minimum wage increase on full time equivalent employees. State which coefficient gives the estimated treatment effect of this policy.

We can write a difference-in-difference estimator in regression form as

$$y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Treatment_i + \beta_3 Post_t \times Treatment_i + u_{it}$$

In this context,

$$FTE_{it} = \beta_0 + \beta_1 Post_t + \beta_2 NJ_i + \beta_3 1993_t \times NJ_i + u_{it}$$

And β_3 tells us the estimated treatment effect in New Jersey in the period following the policy change.

(b) Perform the estimation.

Hint: You will need to create a 'post' dummy variable from the 'year' variable to run this regression. Note that 'state' is already a dummy variable.

```
In [6]: mw <- mutate(mw, post = if_else(year==1993,1,0))

reg1 <- lm(fte ~ post + state + post:state, data = mw)
summary(reg1)

Call:
lm(formula = fte ~ post + state + post:state, data = mw)

Residuals:
    Min       1Q   Median       3Q      Max
-22.174   -6.421   -0.757    4.414   64.243

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   23.673      1.299   18.229  <2e-16 ***
post          -3.499      1.828   -1.914   0.0562 .
state         -2.916      1.444   -2.019   0.0439 *
post:state     1.828      2.025    0.903   0.3671

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.631 on 599 degrees of freedom
(18 observations deleted due to missingness)
Multiple R-squared:  0.01867,    Adjusted R-squared:  0.001375
F-statistic: 3.798 on 3 and 599 DF,  p-value: 0.01023
```

(c) What do you conclude from the results of your estimation (about the differences between NJ and PA, and the effect of the policy change)? Confirm that the results in this part are the same as your estimate in Question 1.2.

Looking at $\hat{\beta}_{NJ}$, we see that New Jersey had significantly lower FTE in fast food stores than Pennsylvania before the policy was enacted: coefficient of -2.916 and p-value of 0.0439. Following the increase in the minimum wage, this gap in FTE employment was reduced -- in other words, employment appears to have increased in New Jersey relative to Pennsylvania by 1.828 FTE. However, the p-value on the coefficient $\hat{\beta}_3$ is 0.3671, so there is no evidence that the minimum wage had any significant effect on employment. In general, with this sample of data our standard error on $\hat{\beta}_3$ is very large, so we cannot reject fairly large positive or negative effects from the increase in the minimum wage.

Question 1.4.

In this question, we will explore the identifying assumptions for the difference-in-differences estimator.

(a) What key assumption do you need to make for your regression in part 1.3 to estimate the causal effect of minimum wage laws?

In order for part 3 to estimate a causal effect of minimum wage increase in minimum wage, we have to assume that without this policy reform, the change in FTE employment for New Jersey pre and post would have been the same as for Pennsylvania. In other words, we assume that the two states would have been on "parallel trends" if not for the policy change, so that the change in FTE employment between New Jersey and Pennsylvania is not due to fundamental differences between the two groups that were already there before the minimum wage increase.

(b) What additional data might you need to provide evidence for this assumption?

While the assumption is fundamentally untestable, we can provide evidence that the trends in employment in the two states were similar before the policy, such that we could have expected them to remain similar in the absence of a minimum wage increase. To provide evidence of "parallel trends" we would need data for the variable fte for both states over several years before the increase in the minimum wage took place.

Question 1.5.

Let's say that we wanted to estimate the effect of minimum wage laws on full-time equivalent employment (fte), but we only had data from New Jersey. Using only data for New Jersey, estimate and interpret the effect of minimum wage laws on full-time employment. Interpret your result, including testing for significance.

Hint: Save the subset of data from New Jersey as a new dataset, and run your regression on that dataset.

```
In [8]: njdata <- mw %>%
  filter(state ==1)

reg2 <- lm(fte ~ post, data = njdata)
summary(reg2)

Call:
lm(formula = fte ~ post, data = njdata)

Residuals:
    Min       1Q   Median       3Q      Max
-21.086   -6.128   -0.421    4.414   64.243

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   20.757      0.609    34.047  <2e-16 ***
post          -1.671      0.836    -1.993   0.0469 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.286 on 490 degrees of freedom
(14 observations deleted due to missingness)
Multiple R-squared:  0.008038,    Adjusted R-squared:  0.006014
F-statistic: 3.971 on 1 and 490 DF,  p-value: 0.04686
```

If we only had data from New Jersey, we would have estimated the change in employment to due to the increase in the minimum wage to be $\bar{Y}_{NJ,post} - \bar{Y}_{NJ,pre} = -1.67$ (this is generated using a regression of fte on post, but we would get the same from a difference in means from part 1.1). This estimate is statistically significant at the 5% level, meaning that we would have (spuriously) concluded that minimum wages decrease employment.

Question 1.6.

In no more than 3 sentences, compare your result from Question 1.4 to your result from Question 1.5. If you draw different conclusions from these results, what might explain this difference, and why might one estimator be preferable?

In Question 1.4, we estimate an increase in full-time employment, although it's not statistically significant. Meanwhile in Question 1.5, we estimate a decrease in full-time employment of a similar magnitude, which is statistically significant. The reason for the discrepancy is that the simple difference does not control for omitted variables that are changing over time across the country (e.g., a nationwide recession). The differences-in-differences estimator accounts for these types of omitted variables by comparing the change in full time employment that occurred in New Jersey to the change in Pennsylvania. Therefore the estimator in question 1.4 is preferable.

Question 1.7.

Consider each of the following statements (that are not necessarily true) and discuss whether it supports, violates, or is irrelevant to assumptions necessary for the DD estimator to provide a valid causal effect in this case. If it violates the assumptions, discuss how it might bias the results.

- New Jersey has long been considered one of the worst places in the US to live because of poverty, crime, and pollution.
- Manufacturing and coal are big parts of Pennsylvania's economy (but not New Jersey's). During the late 1980s and early 1990s, employment in these sectors steadily dropped as the US continued to deindustrialize and open up to trade.
- In 1993, McDonald's reintroduced the McRib, increasing demand for fast-food nationwide.
- In 1993, New Jersey elected a new governor who implemented a sweeping wave of state-level reforms including infrastructure investments, tax cuts, and welfare expansion alongside the minimum wage law.

1. This is **irrelevant**. As long as New Jersey's undesirability relative to Pennsylvania is constant over time (as the statement suggests), this will be picked up by the state dummy and not affect our DD coefficient of interest.

2. This would **violate** the parallel trends assumption. It suggests that employment may have already been falling in Pennsylvania relative to New Jersey, leading to upward bias in our estimates (we would be falsely attributing part of the widening gap in employment between New Jersey and Pennsylvania to the effect of the minimum wage policy).

3. This is **irrelevant**. Any changes in nationwide fast food demand would be picked up by the 1993/post dummy, to the extent that the McRib has similar effects across states. It could create a bias if people are more responsive to the McRib across states (for example, if a larger share of New Jerseyans don't eat pork for dietary reasons, then we might be downward biased because we would have expected employment in Pennsylvania to have been even lower had the McRib not been introduced).

4. This would **violate** the assumption that the only thing changing in the difference between NJ and PA between 1992 and 1993 is the minimum wage policy. We'd be picking up the effect any policies implemented together with the minimum wage -- we'd be upward biased if these policies increase employment overall and downward biased if they decrease it.

Exercise 2: Schoolbus Replacements and Attendance - Panel Regression

Background

In this exercise, we will look at the effect of replacing high-polluting school buses on students' health. Many school districts in California, particularly less wealthy school districts, have school buses that are many decades old. These buses do not have many of the pollution controls that are now standard in vehicles, exposing the students that ride them to high concentrations of pollutants. In 2006, the state of California passed a proposition that allocated funds towards replacing the oldest of these school buses with new models that had adequate pollution controls. We have data for these replacements for the years 2009-2012, with the number of replacements per year more or less increasing over the sample period. This data is combined with attendance data from all school districts in California over the same period to test the impact of reducing pollution exposure through bus replacements on student health. Attendance is used to measure student health because students who are chronically ill are often absent from school. The full dataset is described in detail below.

The dataset Schoolbuses_PS5.dta is an unbalanced panel of 200 school districts for the years 2009-2012, and contains the following variables:

| Variable Name | Description |
|---------------|---|
| district_code | Unique School District Identifier |
| year | Year |
| bus_replace | Number of Buses Replaced |
| attendance | Percent of students in attendance, on average in the year |
| gifted | Number of students in the Gifted Student Program |
| white | Number of White Students |
| college | Number of Students with Parents that Attended College |
| advtdg | Number of Students from Higher Socio-Economic Backgrounds |
| fleet_size | Number of Buses in the District Fleet |
| pupils_trans | Average Number of Students Traveling per Day |
| enrollment | Number of Students Enrolled in the District |

Some summary statistics are provided below

```
In [13]: schooldata <- read.csv("Schoolbuses_PS5.csv")
head(schooldata)

# Summary Stats
busrep <- summarize(schooldata, mean = mean(bus_replace),
  sd = sd(bus_replace),
  min = min(bus_replace),
  max = max(bus_replace))
enroll <- summarize(schooldata, mean = mean(enrollment),
  sd = sd(enrollment),
  min = min(enrollment),
  max = max(enrollment))
fleet <- summarize(schooldata, mean = mean(fleet_size, na.rm = TRUE),
  sd = sd(fleet_size, na.rm = TRUE),
  min = min(fleet_size, na.rm = TRUE),
  max = max(fleet_size, na.rm = TRUE))
ss <- rbind(busrep, enroll, fleet)
sumstats <- cbind(c("Bus Replacements", "Enrollment", "Fleet Size"), ss)
names(sumstats)[1] <- "Variable"
print('Sum Stats')
sumstats
```

A data.frame: 6 x 11

| | district_code | year | bus_replace | attendance | gifted | college | advtdg | fleet_size | pupils_trans | enrollment | |
|---|---------------|-------|-------------|------------|--------|---------|--------|------------|--------------|------------|-------|
| | <int> | <dbl> | <dbl> | <dbl> | <int> | <int> | <int> | <dbl> | <dbl> | <int> | <int> |
| 1 | 261333 | 2012 | 11.750881 | 92.57732 | 0 | 31 | 6 | 25 | 3 | 63.5 | 97 |
| 2 | 461382 | 2010 | 0.000000 | 93.43200 | 17 | 87 | 22 | 36 | 2 | 31.0 | 125 |
| 3 | 461382 | 2011 | 2.060723 | 91.42857 | 14 | 71 | 21 | 28 | 1 | 62.0 | 126 |
| 4 | 461382 | 2012 | 3.942311 | 94.70634 | 13 | 62 | 19 | 25 | 1 | 44.5 | 126 |
| 5 | 461408 | 2010 | 3.942311 | 93.39867 | 18 | 202 | 37 | 134 | 4 | 70.0 | 534 |
| 6 | 461424 | 2012 | 0.000000 | 92.08743 | 776 | 6081 | 1976 | 5670 | 10 | 694.0 | 12364 |

[1] "Sum Stats"

A data.frame: 3 x 5

| | Variable | mean | sd | min | max |
|------------------|------------|-------------|-------|-------------|-------|
| | <chr> | <dbl> | <dbl> | <dbl> | <dbl> |
| Bus Replacements | 2.351899 | 3.315411 | 0 | 12.92825 | |
| Enrollment | 1006.06636 | 1562.073757 | 7 | 12931.00000 | |
| Fleet Size | 5.406863 | 6.550293 | 0 | 95.00000 | |

```
In [10]: #Confirm number of observations
length(unique(schooldata$district_code))

200
```

Question 2.1.

You think that it might be important to control for the year in your regression of attendance on bus replacements. First, generate year dummy variables (Y_{2009} , Y_{2010} , Y_{2011} , Y_{2012}). Next, estimate the following equation for school attendance.

$$attendance_{it} = \beta_0 + \beta_1 bus_replace_{it} + \beta_2 white_{it} + \beta_3 college_{it} + \beta_4 advtdg_{it} + \beta_5 gifted_{it} + \delta_1 + u_{it} \quad (1)$$

(a) Estimate the model and report your results.

```
In [11]: schooldata <- schooldata %>%
  mutate(
    yr_2009 = as.numeric((year == 2009)),
    yr_2010 = as.numeric((year == 2010)),
    yr_2011 = as.numeric((year == 2011)),
    yr_2012 = as.numeric((year == 2012)))

regla <- lm(attendance ~ bus_replace + white + college + advtdg + gifted + yr_2010 + yr_2011 + yr_2012,
  data = schooldata)
summary(regla)

Call:
lm(formula = attendance ~ bus_replace + white + college + advtdg +
  gifted + yr_2010 + yr_2011 + yr_2012, data = schooldata)

Residuals:
    Min       1Q   Median       3Q      Max
-87.881   -0.329    3.258    5.769   19.409

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  90.8219136   1.2919545   70.298  <2e-16 ***
white        -0.0051279   0.0023476   -2.184   0.0294 *
college       0.0005309   0.0075253    0.071   0.9439
advtdg       0.0021730   0.0038630    0.563   0.5740
gifted       0.0119458   0.0098622    1.211   0.2264
yr_2010      -1.5424792   1.7119411   -0.901   0.3680
yr_2011      -0.6899655   1.8375698   -0.375   0.7075
yr_2012      2.3047216   3.0480841    0.756   0.4499

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.62 on 502 degrees of freedom
Multiple R-squared:  0.02724,    Adjusted R-squared:  0.01174
F-statistic: 1.757 on 8 and 502 DF,  p-value: 0.08318
```

(b) Give the meaning (economic interpretation) of β_0 and δ_1

In the presence of year fixed effects, $\hat{\beta}_0$ is the average attendance in 2009 when zero bus replacements have been made, and there are zero white students, students whose parents are in college, advantaged, and gifted students.

δ_1 is the average difference for school attendance in 2010 relative to the omitted year 2009, holding all other factors constant

(c) Interpret $\hat{\beta}_1$. Be sure to mention sign, size and significance and what is being held constant.

An additional bus replacement is associated with a 0.02 percentage point increase in school attendance, holding constant the number of students that are white, gifted, whose parents attended college, and who come from advantaged socio-economic backgrounds as well as overall trends in a given year. This effect is small and not remotely significant ($p > .95$)

(d) Why is the year 2009 dummy excluded?

We can't include the 2009 dummy due to perfect collinearity! Given how we generated the variable, we know that

$$Y_{2009} = 1 - Y_{2010} - Y_{2011} - Y_{2012}$$

So if we tried to include Y_{2009} then R would be unable to compute our regression as we have violated our no multicollinearity assumption.

Question 2.2.

Consider now the following (unobserved) fixed effects model:

$$attendance_{it} = \beta_0 + \beta_1 bus_replace_{it} + \beta_2 white_{it} + \beta_3 college_{it} + \beta_4 advtdg_{it} + \beta_5 gifted_{it} + \delta_i + u_{it} \quad (2)$$

(a) Why are we adding district fixed effects (δ_i)? In other words, what do these fixed effects control for in the regression?

We add the district fixed effects terms, δ_i , to control for all characteristics of a school district that do not change over time (or to be precise, whose effect on attendance does not change over time). This allows us to control for variables (and thereby reduce omitted variable bias and increase the precision of our estimates) without actually having to measure them. The specific point estimate of a particular δ_i can be interpreted as the difference between average attendance of school district i and the omitted school district, after controlling for the other variables in the regression.

(b) Estimate the model and interpret $\hat{\beta}_1$. Be sure to mention sign, size, and significance and what is being held constant.

Hint: Use felm. Remember to use as.factor to turn your fixed effects variables into factors (dummy variables for each category).

```
In [14]: reg2c <- felm(attendance ~ bus_replace + white + college + advtdg + gifted | as.factor(year) + as.factor(district_code),
  data = schooldata)
summary(reg2c)

Call:
felm(formula = attendance ~ bus_replace + white + college + advtdg +
  gifted | as.factor(year) + as.factor(district_code), data = schooldata)

Residuals:
    Min       1Q   Median       3Q      Max
-31.786   -1.148    0.000    1.241   28.522

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
bus_replace  0.4611815   0.2339752    1.971   0.0496 *
white        -0.0053466   0.0063585   -0.841   0.4011
college       0.0170546   0.0140872   -1.211   0.2270
advtdg       0.0008644   0.0032996    0.262   0.7935
gifted       0.0016919   0.0100704    0.168   0.8667

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.658 on 303 degrees of freedom
Multiple R-squared (full model):  0.8987    Adjusted R-squared:  0.8294
Multiple R-squared (proj model):  0.02164    Adjusted R-squared:  -0.6467
F-statistic (full model): 12.98 on 207 and 303 DF, p-value: < 2.2e-16
F-statistic (proj model):  1.34 on 5 and 303 DF, p-value: 0.2471
```

An additional bus replacement is associated with a 0.46 percentage point increase in school attendance, holding constant the share of students that are white, gifted, whose parents attended college, and who come from advantaged socio-economic backgrounds, as well as all unobserved characteristics of a particular district and a particular year. This is significant at the 5% level. There is statistical evidence at 95% confidence that bus replacements increase attendance (conditional on all the above).

(c) Comment specifically on how the size of $\hat{\beta}_1$ changes from model (1) to model (2). Describe your intuition for why it changes in the way that it does, and a possible omitted variable that could explain the difference.

The coefficient moves from $\hat{\beta}_1 = 0.022$ to 0.461. This suggests that our estimate from equation (1) was downward biased when we failed to include school district fixed effects. We might expect that less resource rich school districts will have more bus replacements (because the program targeted the worst buses which are in the poorest districts), but poorer attendance. So if we fail to include these fixed effects it will look like bus replacements don't have much of an effect, but really more bus replacements occur in poorer school district, which have lower attendance for other reasons.

In MLR4 language, you can think of resources, as an omitted variable (that is roughly constant in all years), and if we assume $\text{cov}(\text{resources}, \text{attendance}_i) < 0$ and $\text{cov}(\text{resources}, \text{bus_replace}_i) > 0$, then it makes sense that we would have found a downward bias.

Question 2.3.

What is the MLR 4 assumption for model (2) to be unbiased? Do you think it is likely to hold in this case? Whatever position you take, give your argument.

- MLR4: For each i , the conditional expectation of the idiosyncratic error, u_{it} , given the explanatory variables in all time periods and the fixed effect is zero. $E[u_{it} | x_{it}, \delta_i, \delta_j] = 0$

There were two ways this could have been answered for full marks.

If you argued that MLR4 holds, you would need to say something along these lines: "With the year and school district fixed effects, the only potential sources of omitted variable bias are time-varying factors that are correlated with both the replacement of buses and the attendance rate in schools. One reason to think that there might not be time-varying omitted variables is if the only factor that determines the number of buses that are replaced is the stock of pre-2006 buses, which might not be time-varying."

If you argued that MLR4 would not hold, a complete answer would look something like this: "We might be concerned that there is correlation between the number of buses replaced and any omitted variable that changes over time and affects attendance. A possible omitted variable that could bias our estimates is changes in the demographics of school districts that are both correlated to attendance and bus replacements. For example, if wealthier families move into a district, they may be more likely to have high attendance rate, and also lobby for bus replacements."