Antonella Basso

PHP 2610 Problem Set 3

Due: October 28 by 11:59pm

Instructions: Please upload your answer to the Canvas course page as a pdf file. You can submit your answers in a separate pdf file (please be sure to properly mark the question number to your responses), or you can work on this pdf file, scan it, and upload it to the Canvas course page.

Late or missed assignments: Problem sets and the final report must be turned in online at or before the posted due date. Every one day (24 hours) of delay will result in a ten point (out of 100) downgrade.

Question 1 (30 points)

Please read the abstract from the following paper and then answer questions about it. It is not necessary to read the full paper to answer the questions below:

Raifman, J., Moscoe, E., Austin, S. B., & McConnell, M. (2017). Difference-in-differences analysis of the association between state same-sex marriage policies and adolescent suicide attempts. *JAMA pediatrics, 171*(4), 350-356.

Importance. Suicide is the second leading cause of death among adolescents between the ages of 15 and 24 years. Adolescents who are sexual minorities experience elevated rates of suicide attempts.

Objective. To evaluate the association between state same-sex marriage policies and adolescent suicide attempts.

Design, Setting, and Participants. This study used state-level Youth Risk Behavior Surveillance System (YRBSS) data from January 1, 1999, to December 31, 2015, which are weighted to be representative of each state that has participation in the survey greater than 60%. A difference-in-differences analysis compared changes in suicide attempts among all public high school students before and after implementation of state policies in 32 states permitting same-sex marriage with year-to-year changes in suicide attempts among high school students in 15 states without policies permitting same-sex marriage. Linear regression was used to control for state, age, sex, race/ethnicity, and year, with Taylor series linearized standard errors clustered by state and classroom. In a secondary analysis among students who are sexual minorities, we included an interaction between sexual minority identity and living in a state that had implemented same-sex marriage policies.

Interventions. Implementation of state policies permitting same-sex marriage during the full period of YRBSS data collection.

Main Outcomes and Measures. Self-report of 1 or more suicide attempts within the past 12 months.

Results. Among the 762,678 students (mean [SD] age, 16.0 [1.2] years; 366,063 males and 396,615 females) who participated in the YRBSS between 1999 and 2015, a weighted 8.6% of all high school students and 28.5% of students who identified as sexual minorities reported suicide attempts before implementation of same-sex marriage policies. Same-sex marriage policies were associated with a 0.6–percentage point (95% CI, –1.2 to –0.01 percentage points) reduction in suicide attempts, representing a 7% relative reduction in the proportion of high school students attempting suicide owing to same-sex marriage implementation. The association was concentrated among students who were sexual minorities.

Conclusions and Relevance. State same-sex marriage policies were associated with a reduction in the proportion of high school students reporting suicide attempts, providing empirical evidence for an association between same-sex marriage policies and mental health outcomes.

Consider the following notation

- i: units
- t: time (in years)
- T_0: the number of pre-intervention periods
- A_i: an indicator of living in a state that had implemented same-sex marriage policies
- $\mathbf{X}_i$: conditioning covariates, e.g., state, age, sex, race/ethnicity, and year
- Y_{it}: unit i's self-report of 1 or more suicide attempts within the past 12 months measured at year t
- Y_{it}^1: unit i's potential outcome under the intervention at time t
- Y_{it}^0: unit i's potential outcome under the control at time t

Q1.1. (5 points) What are the treatment and control conditions in this example?

a. Treatment: Attempting to commit suicide once or more in the previous 12 months, Control: Not attempting to commit suicide in the previous 12 months
b. Treatment: Attending public high schools after the implementation of same-sex marriage policies, Control: Attending public high schools before the implementation of same-sex marriage policies
c. Treatment: Living in a state that had implemented same-sex marriage policies, Control: Living in a state that had not implemented same-sex marriage policies
d. Treatment: Living in a state that had experienced elevated rates of suicide attempts, Control: Living in a state that had not experienced elevated rates of suicide attempts
e. None of the above

Answer:

Q1.2 (5 points) Suppose that the consistency assumption holds. Which potential outcome can never be observed (choose all that apply)?

a. Y_{it}^0 with $A_i = 0$ and $t \leq T_0$
b. Y_{it}^1 with $A_i = 0$ and $t > T_0$
c. Y_{it}^0 with $A_i = 1$ and $t \leq T_0$
d. Y_{it}^1 with $A_i = 1$ and $t > T_0$
e. Y_{it}^0 with $A_i = 1$ and $t > T_0$

Answer:

Q1.3. (6 points) Please define the target estimand using the potential outcomes defined above:

Answer:

ATT: E(Y1_t1−Y0_t1|A=1) ,
where t1 indicates the post-intervention period (i.e., for t > T0)

Q1.4. (6 points) Consider the following model with $E(\epsilon_{it}) = 0$. Suppose that D_{it} indicates the treatment status of unit i at time t ($D_{it} = 1$ if intervened and $D_{it} = 0$ otherwise).

$$Y_{it} = \alpha_0 + \alpha_1 I(t > T_0) + \alpha_2 I(A_i = 1) + \alpha_3 D_{it} + \boldsymbol{\beta}^T \mathbf{X}_i + \epsilon_{it}.$$

Please represent the difference in expected outcomes for adolescents living in a state that had implemented same-sex marriage policies but without the intervention before and after the implementation of the policies, using the model coefficients.

Answer: $\alpha 1$

(i) $E[Y|A=1, D=0, t > T0] \;\; \text{->} \;\; \alpha 0 + \alpha 1 + \alpha 2 + \beta$

(ii) $E[Y|A=1, D=0, t \leq T0] \;\; \text{->} \;\; \alpha 0 + \alpha 2 + \beta$

(i) - (ii) = $E[Y0\_t1 - Y0\_t0|A=1] = \alpha 1$

Q1.5. (8 points) Consider the following model with $E(\epsilon_{it}) = 0$ and time-varying coefficients $\boldsymbol{\beta}_t$. Suppose that D_{it} indicates the treatment status of unit i at time t ($D_{it} = 1$ if intervened and $D_{it} = 0$ otherwise).

$$Y_{it} = \alpha_0 + \alpha_1 I(t > T_0) + \alpha_2 I(A_i = 1) + \alpha_3 D_{it} + \boldsymbol{\beta}_t^T \mathbf{X}_i + \epsilon_{it}.$$

Please represent the difference in expected outcomes (i) for adolescents living in a state that had implemented same-sex marriage policies without the intervention before and after the implementation of the policies, conditional on the covariates $\mathbf{X}_i$, using the model coefficients. (ii) Derive the same trends for adolescents living in a state that had not implemented same-sex marriage policies. (iii) Compare the two quantities from (i) and (ii).

Answer:

(i.a) $E[Y|A=1, D=0, t>T0] \;\; \text{->} \;\; \alpha 0 + \alpha 1 + \alpha 2 + \beta\, E[X|A=1, D=0, t>T0]$

(i.b) $E[Y|A=1, D=0, t \leq T0] \;\; \text{->} \;\; \alpha 0 + \alpha 2 + \beta\, E[X|A=1, D=0, t \leq T0]$

(i.a) - (i.b) = $E[Y0\_t1 - Y0\_t0|A=1] = \alpha 1 + ((\beta 0\_t1 - \beta 0\_t0)\, E[X|A=1])$

(ii.a) $E[Y|A=0, D=0, t>T0] \;\; \text{->} \;\; \alpha 0 + \alpha 1 + \beta\, E[X|A=0, D=0, t>T0]$

(ii.b) $E[Y|A=0, D=0, t \leq T0] \;\; \text{->} \;\; \alpha 0 + \beta\, E[X|A=0, D=0, t \leq T0]$

(ii.a) - (ii.b) = $E[Y0\_t1 - Y0\_t0|A=0] = \alpha 1 + ((\beta 0\_t1 - \beta 0\_t0)\, E[X|A=0])$

(i) - (ii) = $[\alpha 1 + ((\beta 0\_t1 - \beta 0\_t0)\, E[X|A=1])] - [\alpha 1 + ((\beta 0\_t1 - \beta 0\_t0)\, E[X|A=0])]$
$= [(\beta 0\_t1 - \beta 0\_t0)\, E[X|A=1]] - [(\beta 0\_t1 - \beta 0\_t0)\, E[X|A=0]] \neq 0$

Quantities from (i) and (ii) cannot be assumed to be equal in the presence of time-varying covariates, as this will, in many cases, violate the parallel trends assumption, thereby inducing bias on our estimate of the ATT. However, there are cases in which the parallel trends assumption may hold, conditionally, when dealing with such variables. For example, in the case that changes in time-varying covariates are unrelated to the intervention (i.e., $E[X|A=1]=E[X|A=0]$).

Question 2 (30 points)

Please read the abstract from the following paper and then answer questions about it. It is not necessary to read the full paper to answer the questions below:

Rudolph, K. E., Stuart, E. A., Vernick, J. S., & Webster, D. W. (2015). Association between Connecticut's permit-to-purchase handgun law and homicides. *American journal of public health, 105*(8), e49-e54.

Objectives. We sought to estimate the effect of Connecticut's implementation of a handgun permit-to-purchase law in October 1995 on subsequent homicides.

Methods. Using the synthetic control method, we compared Connecticut's homicide rates after the law's implementation to rates we would have expected had the law not been implemented. To estimate the counterfactual, we used longitudinal data from a weighted combination of comparison states identified based on the ability of their prelaw homicide trends and covariates to predict prelaw homicide trends in Connecticut.

Results. We estimated that the law was associated with a 40% reduction in Connecticut's firearm homicide rates during the first 10 years that the law was in place. By contrast, there was no evidence for a reduction in nonfirearm homicides.

Conclusions. Consistent with prior research, this study demonstrated that Connecticut's handgun permit-to-purchase law was associated with a subsequent reduction in homicide rates. As would be expected if the law drove the reduction, the policy's effects were only evident for homicides committed with firearms. (Am J Public Health. 2015;105:e49-e54. doi:10.2105/AJPH.2015. 302703)

Consider the following notation

- i: units
- t: time (in years), $t = 1, 2, \ldots, T$. Let $T_0 (< T)$ denote the number of pre-intervention periods
- A_i: an indicator of unit i being in the intervention group

- Y_{it}: firearm-specific homicide rates of unit i at time t

---

Q2.1. (5 points) What are the units of this analysis?

a. Residents living in Connecticut
b. Handgun purchasers in Connecticut
c. States in the US
d. Cities in the US
e. None of the above

Answer:

Q2.2. (5 points) People who lived near the border of Connecticut and other states who can cross the border to purchase a handgun may cause a problem in terms of which assumption?

a. Consistency
b. Positivity
c. The Stable Unit Treatment Value Assumption
d. Unconfounded treatment assignment
e. Exclusion restriction
f. None of the above

Answer:

Q2.3. (5 points) What would be the donor pools in this study?

a. Residents living in the US other than Connecticut
b. Residents living in Connecticut who did not purchase a handgun
c. Other states with no permit-to-purchase handgun law
d. Other states with population characteristics that are comparable to Connecticut's
e. Other states than Connecticut with permit-to-purchase handgun law and with similar population characteristics with those of Connecticut

Answer:

Q2.4. (7 points) Let Z_t indicate whether time t is in the post-intervention period. Suppose that we fit the following model instead of using the synthetic control method:

$$Y_{it} = \beta_0 + \beta_1 t + \beta_2 A_i + \beta_3 Z_t + \beta_4(t \times A_i) + \beta_5(A_i \times Z_t) + \beta_6(t \times Z_t) + \beta_7(t \times A_i \times Z_t) + \epsilon_{it},$$

where ϵ_{it} is a residual term with mean zero.

Please represent the Connecticut's post-intervention's expected outcome at t ($t > T_0$) in a linear form using the coefficients.

Answer: **ß0 + tß1 + ß2 + ß3 + tß4 + ß5 + tß6 + tß7 = ß0 + ß2 + ß3 + ß5 + t(ß1 + ß4 + ß6 + ß7)**

Q2.5. (8 points) Using the regression model in the above, represent the pre-intervention (desirable) expected outcome of the synthetic control at t ($t \leq T_0$):

Answer: **ß0 + tß1**

Question 3 (40 points)

This assignment uses observational data to examine the smoking ban in public places on the acute coronary episodes . The data (`sicily.csv`) is available Canvas/Files/Data. For more information, please refer to Bernal et al. (2017) (https://doi.org/10.1093/ije/dyw098): Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International journal of epidemiology, 46*(1), 348-355.

The variables in the dataset are:

- `year`: year
- `month`: month
- `time` : elapsed time since the start of the study
- `aces` : count of acute coronary episodes (ACE) in Sicily per month (the outcome)
- `smokban` : smoking ban (the intervention) coded 0 before intervention, 1 after
- `pop` : the population of Sicily (in 10000s)
- `stdpop` : age standardised population

For your reference for Poisson regression for rates, this page (https://rpubs.com/kaz_yos/poisson) may be useful.

```
# load the data
data <- read.csv("sicily.csv")

# compute the standardized rates
data$rate <- with(data, aces/stdpop*10^5)

# fit a Poisson with the standardised population as an offset
model1 <- glm(aces ~ offset(log(stdpop)) + smokban + time, family=poisson,
   data)
```

Q3.1. (7 points) Based on the results from '`model1`', please calculate the predicted count of ACE for Year 2005 Month 5 without the intervention (round to 2 decimal places).

Answer: **913.40** **`` `predict(model1, type="response", data2)[41]` ``**

Q3.2. (7 points) Based on the results from '`model1`', please calculate the predicted count of ACE for Year 2005 Month 5 under the intervention (round to 2 decimal places).

Answer: **816.92** **`` `predict(model1, type="response", data)[41]` ``**

Q3.3. (4 points) Please derive the estimate associated with the intervention effect and its 95% confidence interval using the following command. Round to 2 decimal places and for the 95% confidence interval, answer in the form (L, U), where L and U indicate the lower and upper bounds, respectively.

```
coefficients(model1)
confint(model1)
```

Answer: **Estimated "smokban" effect ≈ -0.11, with a 95% CI of (L, U) ≈ (-0.15, -0.08).**

Q3.4. (6 points) Please describe the relationships between the differences in your answers in Q.3.1. and Q.3.2. and your estimate in Q.3.3

Answer: **(see below)**

Q3.5. (8 points) Please consider the following model, now with an interaction term

```
# fit a Poisson with the standardised population as an offset
model2 <- glm(aces ~ offset(log(stdpop)) + smokban*time,
              family=poisson, data)
```

Based on the results from 'model2', please calculate the predicted difference in log(count of ACE) for Year 2005 Month 5 between the intervention and control status (round to 3 decimal places).

Answer: **-0.112** **predicted log(ACE count) under intervention status - predicted log(ACE count) under control status** → `` `predict(model2, data)[41] - predict(model2, data2)[41]` ``

Q3.6. (8 points) Based on the results from 'model2', please calculate the predicted difference in log(count of ACE) for Year 2006 Month 3 between the intervention and control status (round to 3 decimal places).

Answer: **-0.107** **predicted log(ACE count) under intervention status - predicted log(ACE count) under control status** → `` `predict(model2, data)[54] - predict(model2, data2)[54]` ``

Q3.4.

Given `model1`, the predicted value from Q3.1, which assumes no intervention (unobserved counterfactual), is approximately exp(-0.11) ≈ 0.89 times what we observe in the predicted outcome from Q3.2 with intervention. That is, Q3.1 gives an estimate of what we would have observed (i.e., what the model would have predicted) had there not been an intervention. Namely, we would have seen a higher ACE count had the smoke ban not taken place as the intervention has a negative effect on the outcome. Using the effect estimate and its 95% CI to alternate between the two quantities (Q3.1 to Q3.2 and vise versa) confirms this. That is, with quantities from Q3.3, `effect_est` ≈ (-0.15, -0.11, -0.08), we see that:

```
`pred_1/exp(effect_est)` ≈ (944.75, 913.40, 883.06) and
`pred_0*exp(effect_est)` ≈ (789.82, 816.92, 844.99),
```

give the quantities and CI's from Q3.1 and Q3.2 in the form (L, effect, U), respectively.