PHS 2000B: Difference-in-differences and interrupted time series

Requirements for Difference-in-Differences and Interrupted Time Series

Difference-in-Differences (DD) and Interrupted Time Series (ITS) are popular methods of estimating causal effects in economics. DD is a design which is usually applied when treatment is administered at a group level. ITS, on the other hand, is usually used to evaluate the effect of interventions that have been implemented at a defined point in time.

Both methods generally have specific data and contextual requirements. In terms of data, DD and ITS require data that have some form of time or cohort dimension to it. DD can be applied when we have: a) cross-sectional data which has some sort of group (for example, county of birth) and time dimension to it (for example, year of birth); b) repeated cross-sections, i.e., data consisting of two or more (random) samples taken at different time points from the same population but not the same people; or, c) panel data, i.e., data consisting of a sample from which information is collected over multiple time points.

ITS, on the other hand, can only be applied when we have access to time series data. Time series data refers to information collected from the same unit with high frequency over a certain length of time. A heuristic to distinguish between panel and time series data is that panel datasets usually have large N (sample size) but small T (number of time points in which data are collected) while time series data usually have small N (often just N = 1) but large T.

In terms of context, for both designs, we need to be able to clearly distinguish between a pre-treatment period and a post-treatment period. In the case of ITS, it is also helpful to have outcomes that we expect will change relatively quickly upon the implementation of an intervention. In the case of a Controlled Interrupted Time Series (CITS), we also need an outcome time series for a viable comparison group.

Causal estimand

Because we're introducing a time component into our analysis, it's worth explicitly defining the notation we will be using in this lab. A will refer to a treatment variable and for now we will assume it to be binary such that $A = 1$ implies treatment assignment. Y refers to an arbitrary outcome. Y_{it}^{a} refers to the potential outcome for individual i in time period t when treatment A is set to a, possibly contrary to fact. So, for example, $Y_{i0}^{a=1}$ refers to the outcome for individual i in time period 0 when the treatment A is set to 1, possibly contrary to fact.

Difference-in-differences

The canonical DD design has two time periods $t = 0, 1$ where $t = 0$ refers to the pre-treatment period and $t = 1$ equals the post-treatment period. The treatment effect or the causal estimand of interest (δ_{DD}) is:

$$\delta_{DD} \equiv E[Y_{i1}^{a=1} - Y_{i1}^{a=0} | A = 1]$$

In other words, the causal estimand of interest is the average difference in outcomes under treatment ($a = 1$) and no treatment ($a = 0$) in the treated group(s) ($A = 1$) in the post-treatment time period $t = 1$.

Interrupted time series

The causal estimand of interest in an ITS study depends on the proposed impact model. An impact model is a researcher's hypothesis about the nature of the impact of the treatment on the outcome. It involves answering questions such as:

1. When will the intervention impact the outcome? Will the impact be immediate? Will it be delayed? Will there be some sort of anticipatory effect?
2. Will the intervention lead to a change in the level of the outcome?
3. Will the intervention lead to a change in the rate of change of the outcome?
4. Will the impact of the intervention be constant over time?

For now, let's assume that the intercept and slope shifts are immediate. Like in the multiple time period DD set-up, let $T = \{0, ..., T_0, ...T\}$. For $t \in T$, assume that $t \leq T_0$ represents the pre-treatment period and $t > T_0$ represents the post-treatment period. Given this,

1. An intecept shift can be written as,

$$\delta_{ITS,intercept} \equiv E[Y_{iT_0}^{a=1} - Y_{iT_0}^{a=0}]$$

2. Writing out a change in slope using potential outcomes notation requires new notation. Let $f(t) = E[Y_{it}^{a}]$, i.e., $f(t)$ is a function describing the potential outcome under treatment $A = a$ at time $T = t$. Assume that $f(t)$ is continuously differentiable and that $f_{t \leq T_0}(t)$ represents the potential outcome function in the pre-treatment period and $f_{t > T_0}(t)$ represents the potential outcome function in the post-treatment period.[1] Finally, let $f'(t)$ represent the first derivative of the function $f(t)$. Then, a change in slope can be written as,

$$\delta_{ITS,slope} \equiv f'_{t > T_0}(t) - f'_{t \leq T_0}(t)$$

Controlled Interrupted Time Series

As in an ITS, there are two possible causal estimands of interest in a CITS: a level shift and a slope shift. Both level and slope shifts can be written out in potential outcomes notation in the same way as in an ITS with one difference: now, we must define these impacts specifically for the treated group.

What type of causal effect are we estimating?

Note, in both the DD and ITS designs, we are estimating an Average Treatment Effect on the Treated (ATT). The fact that we are estimating an ATT may limit the generalizability of our results.

With DD in particular, it is usually not the case that we are estimating the impact of interventions that have random assignment; rather, we are usually estimating the impact of programs that are targeted at specific individuals or areas. Because of this, the treated and control groups may, in fact, be quite different from one another, making the $ATT \neq ATU$ (ATU = Average Treatment Effect on the Untreated). Since the Average Treatment Effect (ATE) is a weighted average of the ATT and ATU, when $ATT \neq ATU$, $ATT \neq ATE$ as well.

[1] In other words, for a binary treatment $A = 0, 1$ where $A = 1$ means treatment receipt, $f_{t \leq T_0}(t) = E[Y_{it}^{a=0}]$ and $f_{t > T_0}(t) = E[Y_{it}^{a=1}]$.

Identifying assumptions

Difference-in-differences

In a DD study, we are interested in identifying a *level shift*.[2] We identify this level shift in a DD analysis by making assumptions about *outcome trends*.

Specifically, in the context of the canonical DD with $t = \{0, 1\}$, we assume:

$$E[Y_{i,1}^{a=0} - Y_{i,0}^{a=0}|A = 1] = E[Y_{i,1}^{a=0} - Y_{i,0}^{a=0}|A = 0]$$

The left hand side of the equation above represents the average trend in the outcome in the treated group(s) in the absence of treatment (note the superscripted $a = 0$). The right hand side of the same equation represents the average trend in the outcome in the control group(s) in the absence of treatment. Thus, in a canonical DD, we assume that the trend in the outcome in the treated group under no treatment would have evolved in the same way between $t = 0$ and $t = 1$ as the trend in the outcome in the control group between the same time periods. This is known as the **common trends assumption**.The common trends assumption is **not testable using our data**. Note that the left hand side of the equation above includes a quantity that we cannot observe, namely $E[Y_{i,1}^{a=0}|A = 1]$.

In the medical literature, you may find the DD identifying assumptions being expressed a bit differently. Specifically, researchers may argue that the identifying assumptions for a DD analysis are:

1. **Having parallel pre-trends:** This means that the trend in the outcome in the pre-treatment period in the treated and control groups are parallel. Researchers, if they have data available, provide statistical evidence to show that the trend in the outcome in the treated group in the pre-treatment period evolves in the same way as the control group over the same time period.
2. **Common shocks:** This means that if anything affecting the outcome occurs in the post-treatment period, it affects both the treated and control groups equally.

Note that in a canonical DD design with only one pre- and one post-treatment periods, we cannot provide statistical evidence for parallel pre-trends. Therefore, you will often find more qualitative arguments for the plausibility of parallel pre-trends in those studies. Furthermore, note that the abovementioned "assumptions" are neither necessary nor sufficient to ensure the actual common trends assumption. The problem largely lies with the parallel pre-trends assumption: logically, it is neither necessary nor sufficient for trends in the outcome in the absence of treatment in the post-treatment period in the absence of treatment to be parallel in the treated and control groups just because outcome trends in the pre-treatment period are parallel.

Interrupted time series

Identifying the treatment effect in an ITS design relies on the assumption that absent of any systematic change, how the outcome changed in the past is the best predictor for how it will change in the future.

We can express the identifying assumption of an ITS as follows:

1. The timing of the treatment is as good as random

This assumption allows us to argue that the treatment does not coincide with other interventions that could plausibly lead to an intercept shift and that the trend in the outcome in the pre-treatment period can be extrapolated to the post-treatment period as the counterfactual trend under no treatment.

Controlled Interrupted Time Series

Suppose we do not posit an anticipation effect or a delayed effect. Furthermore, suppose we propose an impact model with both an intercept shift and a slope change. Then, we can write the identifying assumptions as:

[2]Recall that all causal estimands for DD are defined for the same post-treatment time period or, in the case of the average ATT in the multi time-period DD, as an average over all time periods in the post-treatment period.

1. The timing of the treatment is as good as random
2. The shift in the intercept in the control group at $t = T_0$ would be equal to the shift in the intercept in the treated group at the same time point had the treated group not been treated
3. The change in the trend in the outcome in the control group in between the pre- and post-treatment periods would be the change in the trend in the outcome in the treated group over the same time period had the treated group not been treated

Violations of the identifying assumptions

Difference-in-differences

The parallel trends assumption can be violated in several ways:

1. Selection into treatment based on factors whose impact on the outcome is also time-varying. For a formal presentation of this idea, please see **Appendix 3**.
2. Composition of treated and control groups, defined based on factors that affect both selection into treatment and the outcome, changes between the pre- and post-treatment periods. This issue may be of particular concern when conducting DD analyses using repeated cross-sections. For a formal presentation of this idea, please see **Appendix 3**.
3. Economists argue that the parallel trends assumption may be more valid over a *shorter time horizon*. As such, the choice of the study period may influence whether or not the parallel trends assumption is valid.
4. If events that influence the outcome occur differntially across treatment and control groups (i.e., *uncommon shocks*), then the parallel trends assumption will be violated and treatment effect estimates will be biased.
5. If, for some reason, the outcome is measured differently in the treated and control groups, and if the difference in the way the outcome is measured differentially impacts the trend of the outcome, then the parallel trends assumption will be violated.

Interrupted Time Series

ITS and CITS identifying assumptions may be violated for the following reasons:

1. Unmeasured events that affect the outcome co-occur with the treatment of interest. This may not allow us to identify the intercept shift or the slope change. Within the CITS framework, violations of the identifying assumption can occur if (unmeasured) events that affect the outcome occur differentially across the treated and control groups.

2, The method by which the outcome is collected or measured changes at the time of the intervention. If, previously, the outcome was under-estimated (say, because there wasn't a lot of attention to it) and now it is estimated correctly, we would capture this change in measurement as part of our estimated treatment effect.

Estimation

Difference-in-differences

Canonical difference-in-differences

Suppose we have 2 groups (denoted by S), and 2 time periods (denoted by T). We may have many individuals within these groups, denoted by i. The canonical DD or the 2x2 DD model is:

$$Y_{it} = \alpha + \beta T_i + \gamma Post_t + \delta T_i * Post_t + u_{it}$$

Where,

- Y_{ist} is the outcome of individual i in group s at time t.
- T_s is a variable which equals 1 if the group in which an individual lives is a treatment group and 0 otherwise.
- $Post_t$ is a variable which equals 1 if an observation in the control group is measured in the post-treatment period, i.e., the period of time after treatment is implemented.
- $T_s * Post_t$ is a variable which equals 1 if an observation in the treatment group is measured in the post-treatment period.

We can interpret the coefficients as follows:

- α is the average value/level of the outcome in the control group(s) in the "pre" period
- β is the difference in the average value/level of the outcome between the treated and control groups in the "pre" period
- γ is the change in the average value/level of the outcome in the control group between the "pre" and "post" periods
- δ is the difference-in-differences estimate, *i.e.,* the difference between: 1) the change in the average value/level of the outcome in the treated group between the "pre" and "post" periods; and 2) the change in the average value/level of the outcome in the control group between the "pre" and "post" periods

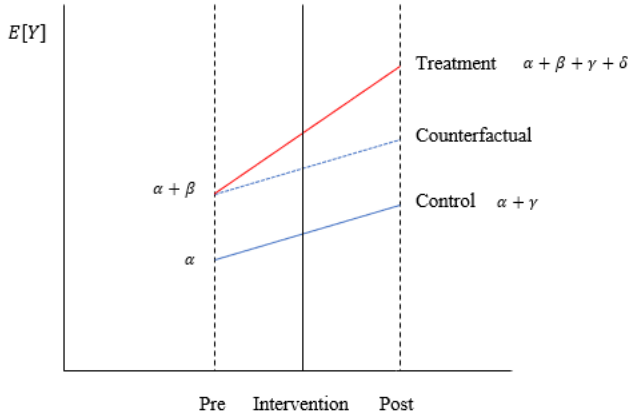


Figure 1: 2x2 DD

Difference-in-differences with group fixed effects

Suppose we have S groups, an arbitrary number of time periods (denoted by T), and individuals i who are within group S. We can now estimate a DD model with group fixed effects as follows:

$$Y_{ist} = \alpha + \sum_{s=2}^{S} \beta_s + \gamma Post_t + \delta T_i * Post_t + \epsilon_{ist}$$

We can interpret the coefficients as follows:

- α is the average value/level of the outcome in the pre-treatment period in the reference group S
- β_s represents group-specific coefficient. For any arbitrary group S, β_s represents the difference in the average value/level of the outcome in the pre-treatment period between group S and the reference group
- γ represents that average change in the value/level of the outcome in the post-treatment period among all groups that did not receive the treatment

- δ is the difference-in-differences estimate. It represents the additional change in the value/level of the outcome in the post-treatment period among all groups that received treatment relative to the average value/level of the outcome in the post-treatment period among groups that did not receive treatment (i.e., relative to γ)

A few points:

1. We do not include a "main effect" for T_i in the model because T_i is a function of the group. From an estimation standpoint, this means that we will not be able to estimate this main effect because T_i is collinear with the group fixed effects.
2. Why do we include group fixed effects in the first place? Recall that group fixed effects allow us to account for group-specific, time-invariant sources of confounding. Therefore, we may want to include these fixed effects to explicitly account for these sources of confounding.
3. We assume that the effect of the treatment on the treated group is homogenous in the post-treatment period.

Finally, in the Economics literature, you will usually see this model written out without an explicit intercept term:

$$Y_{ist} = \sum_{s=1}^{S} \beta_s + \gamma Post_t + \delta T_i * Post_t + \epsilon_{ist}$$

In this model, δ remains the difference-in-differences estimate.

Difference-in-differences with group and time fixed effects

Suppose we have S groups, some of which receive treatment. Further suppose that we have measured the outcome in these groups at across T time periods, some of which are in the pre-treatment period. Finally, suppose that we have measured the outcome at the individual level. We can now estimate a DD model with group and time fixed effects as follows:

$$Y_{ist} = \alpha + \sum_{s=2}^{S} \beta_s + \sum_{t=2}^{T} \gamma_t + \delta T_i * Post_t + \epsilon_{ist}$$

We can interpret the coefficients as follows:

- α is the average value/level of the outcome in the reference time period for the reference group
- β_s is a group-specific coefficient which represents the difference in the average value/level of the outcome in an arbitrary group s relative to the reference group in the reference time period
- γ_t is a time-specific coefficient which represents the difference in the average value/level of the outcome in an arbitrary time period t relative to the reference time period
- δ is the difference-in-differences estimate. It represents the additional change in the value/level of the outcome in the post-treatment period among the groups that received treatment.

A few points:

1. In this model, we do not include a T_i main effect and a $Post_t$ main effect. Again, this is because if we included T_i as a main effect, we would not be able to estimate it due to collinearity with the group fixed effects. Similarly, if we included $Post_t$ as a main effect, we would not be able to estimate it due to collinearity with the time fixed effects.
2. In the previous sub-section, we discussed why we may want to include group fixed effects. Why might we want to include time fixed effects? Time fixed effects allow us to model any trends in the outcome in a very flexible manner.

3.We assume that the effect of the treatment on the treated is homogenous across all time periods in the post-treatment period. This may be a problematic assumption because the effect of the program may vary over time for various reasons (for example, it may be that over time, an organization implementing a specific program becomes more efficient at implementing it, leading to better outcomes). As Zeldow and Hatfield (2019) argue, this should not discourage you from using this model; rather, it should discourage you from automatically using this model. Unlike the two DD models we discussed earlier, we can, in theory, try to model heterogenous effects over time by interacting the treatment indicator T_i with each time fixed effect in the post-treatment period.

4. What if the "Post" period begins at different times for different groups? A recent result by Goodman-Bacon (2019) shows that in such a case, δ represents a weighted average of all possible 2x2 comparisons in our data. For more on this, please see **Appendix 2**.

Finally, in the Economics literature, you will usually see this model written out without an explicit intercept term:

$$Y_{ist} = \sum_{s=1}^{S} \beta_s + \sum_{t=1}^{T} \gamma_t + \delta T_i * Post_t + \epsilon_{ist}$$

In this model, δ remains the difference-in-differences estimate.

Confounding in difference-in-differences

In previous classes, we've estimated causal effects by trying to ensure that our treated and control groups are exchangeable at baseline. Sometimes, as in a randomized trial, we expect treated and control groups to be marginally exchangeable. Other times, we hypothesize that exchangeability is conditional on a set of covariates because of which we control for them in our regression models. The covariates that we control for to create conditional exchangeability are known as confounders.

In the case of a DD study, the notion of confounding - and consequently, what counts as a confounder - is fundamentally different. Recall that the identifying assumption in a DD study is *not* about treated and control groups being (conditionally) exchangeable at baseline but rather about trends in the outcome. Zeldow and Hatfield (2019) argue that in a DD study, a random variable C is a confounder if it satisfies the following two criteria:

1. C is associated with the treatment
2. The relationship between C and the outcome is time-varying **or** the distribution of the covariate across the treated and control groups changes differentially over time

For more on this, please see **Appendix 3**.

Interrupted time series

As stated before, any ITS analysis begins by stipulating the impact model. The following sub-sections will present models under an impact model which stipulates an immediate change in the level as well as a change in the rate of change of the outcome. We also assume that the time variable in the following sub-sections has been centered at the start of the intervention period.

Interrupted time series with immediate intercept shift and slope change

Let i index the unit of observation and t the time period. We can model an immediate intercept and slope change ITS model as follows:

$$Y_{it} = \alpha + \beta time_t + \gamma Post_i + \delta Post_i * time_t + \epsilon_{it}$$

In this model,

- $time_t$: is a continuous variable that measures time
- $Post_i$: is an indicator variable which equals 1 in the "post" period and 0 otherwise

We can interpret these coeffecients as:

- α: Average value/level of the outcome right before treatment is implemented (i.e., $time = 0$ and $Post = 0$)
- β: Slope of the outcome (alternatively, time trend) in the "pre" period
- γ: Change in the average value/level of the outcome (alternatively, intercept shift) at the time of treatment implementation (i.e., $time = 0$ and $Post = 1$)
- δ: Change in the slope of the outcome (alternatively, change in the time trend) in the "post" period

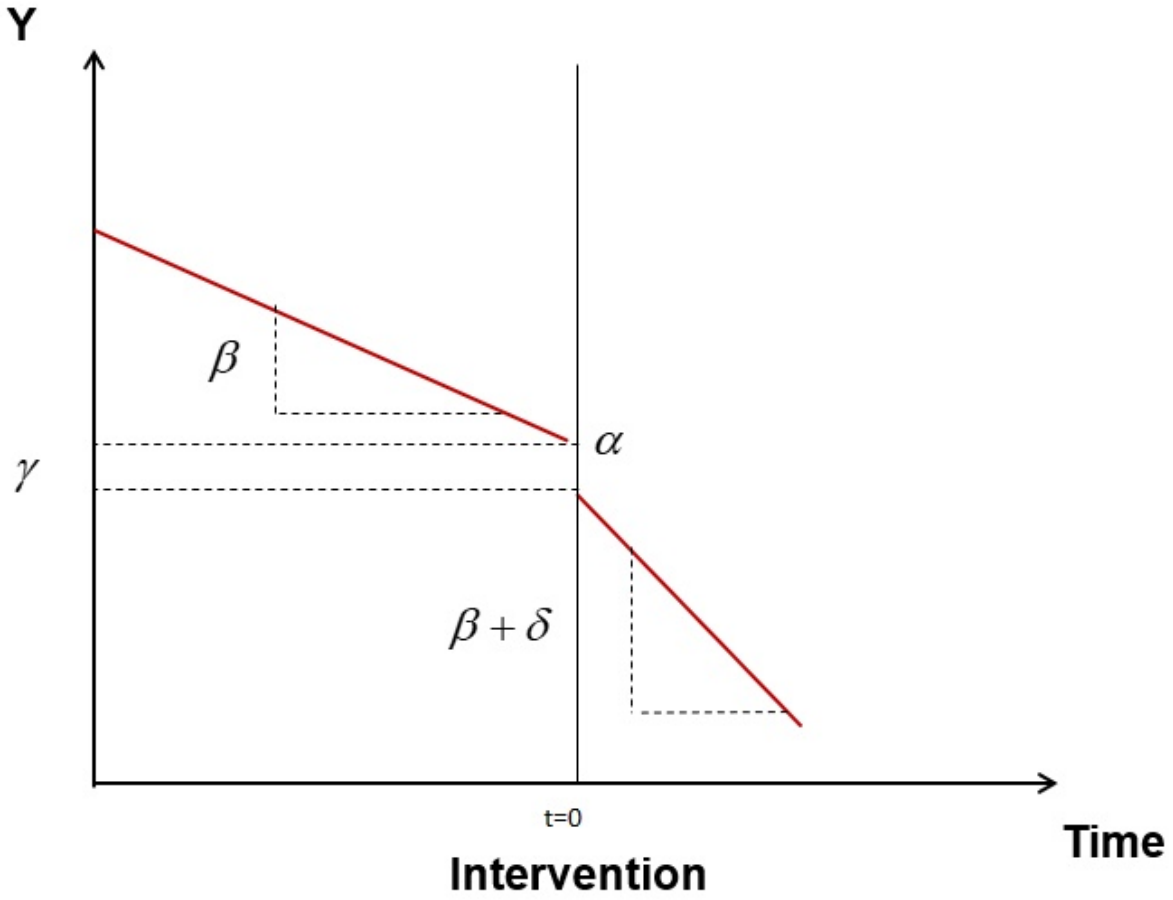


Figure 2: ITS with immediate intercept shift and slope change

Controlled interrupted time series with immediate intercept shift and slope change

Suppose we have time series data on the outcome for both the treated group and a control group. We can now model a CITS with an immediate intercept shift and slope change as follows:

$$Y_{it} = \lambda_0 + \lambda_1 time_t + \lambda_2 Post_i + \lambda_3 time_t * Post_i +$$
$$\alpha treated_i + \beta time_t * treated_i + \delta Post_i * treated_i + \gamma time_t * Post_i * treated_i + \epsilon_{it}$$

In this model,

- $time_t$: is a continuous variable that measures time
- $Post_i$: is an indicator variable which equals 1 in the "post" period and 0 otherwise
- $treated_i$: is an indicator variable which equals 1 if the observation is in the treated group and 0 otherwise

The λ coefficients are all related to the control group.

- λ_0: The value/level of the outcome in the control group right before the start of treatment
- λ_1: The time trend in the control group in the pre-treatment period
- λ_2: The intercept shift in the control group at the start of the treatment period
- λ_3: The change in the time trend in the control group in the post-treatment period

The α, β, δ, and λ coefficients are related to the treated group.

- α: The difference in the value/level of the outcome between the treated and control group right before the start of treatment
- β: The difference in the time trend between the treated and control group in the pre-treatment period
- δ: The additional intercept shift in the treated group relative to the control group at the start of the treatment period
- λ: The additional change in the time trend in the treated group relative to the control group in the post-treatment period

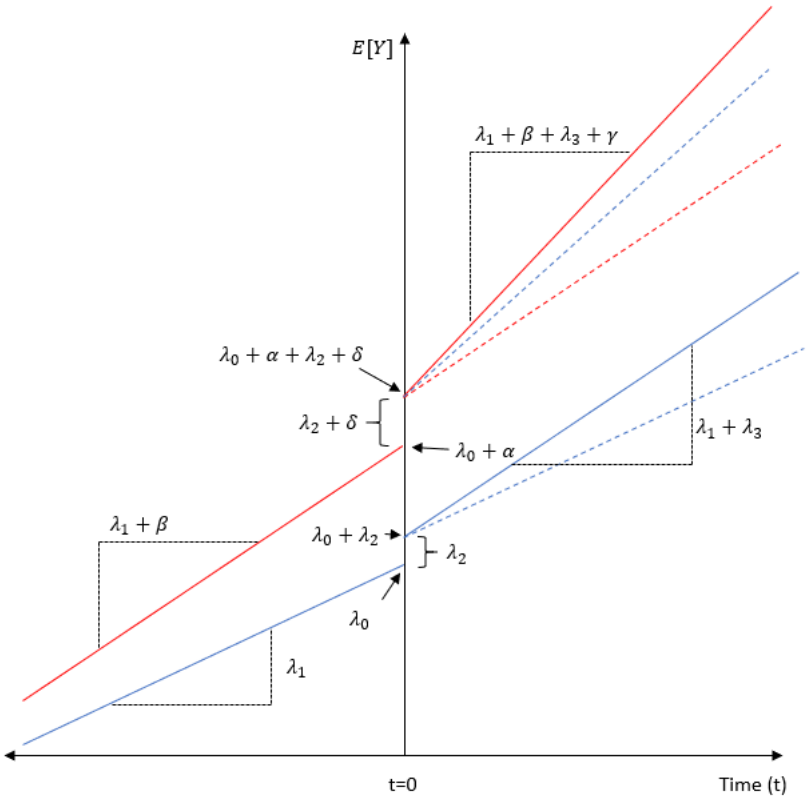


Figure 3: CITS with immediate intercept shift and slope change

Interrupted time series accounting for seasonality in the outcome

In the ITS and CITS models in the previous two sub-sections, we assumed that the time trend can be adequately described in a linear manner. In most cases, however, we will be in a situation where we need to explicitly model seasonal variation in the data. One reason to account for seasonality in our estimation

procedure is to ensure that we model autocorrelated outcomes. A second reason is that sometimes the length of the time series before and after the intervention may be different in length. In this case, by not accounting explicitly for seasonal variation in the outcome, we may get bias in our effect estimates simply because we did not correctly estimate the time trend pre-intervention or the average value of the outcome at the start of the intervention.

The most flexible manner to account for seasonal trends is to estimate time fixed effects defined at the most granular unit of time. However, this is usually not feasible in an ITS/CITS analysis because we have one observation at one time point for the treated and control groups.

Instead, what we can do is model the time trend in the outcome using some flexible function of time. We could write out the model as follows:

$$Y_{it} = \alpha + \beta_f f(time_t) + \beta time_t + \gamma Post_i + \delta Post_i * time_t + \epsilon_{it}$$

In this model,

- $f(time_t)$: is a flexible function of a continuous variable that measures time
- $Post_i$: is an indicator variable which equals 1 in the "post" period and 0 otherwise

The choice of the function does not affect the interpretation of the coefficients. Furthermore, the choice of the function depends on the data itself. For instance, consider the following figure from Bernal et. al. (2016):

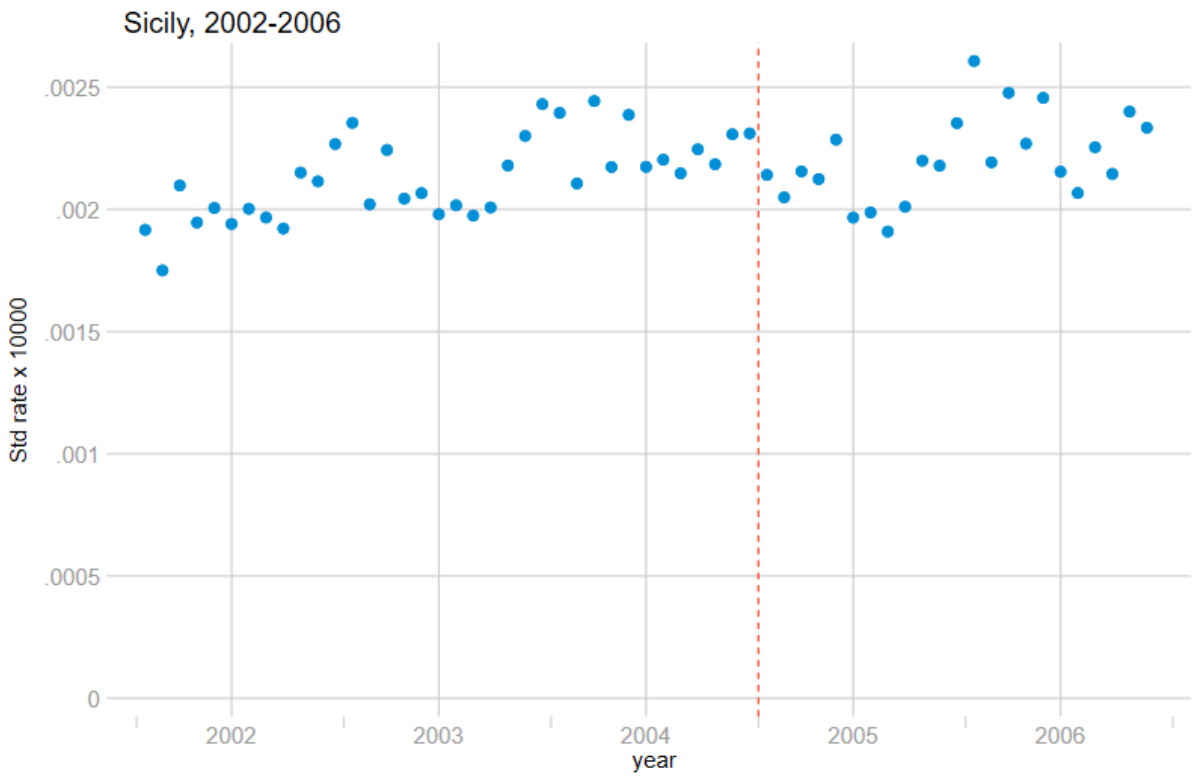


Figure 4: Sicily smoking ban ITS analysis

We could model the time trend in the outcome using some sort of sin or cosine function of time.

Inference

Difference-in-differences

Bertrand, Duflo, and Mullainathan (2004) showed that many existing DD studies demonstrated high levels of intracluster correlation and not accounting for this in estimating standard errors led to incorrect statistical inferences. The topic of how best to estimate standard errors in a DD design has been subject to much research. Wing et. al. (2018) find that although there is no consensus on this issue, there are some consistent themes across a number of papers:

1) Aggregate the outcome to the level at which treatment is assigned
2) If there are a large number of clusters (>50), use the cluster-robust variance estimator
3) If there are fewer clusters, then:

a) Use bootstrap methods (for example, the cluster Wild bootstrap)
b) Use some other approach (such as creating cluster-level randomization distributions)

Interrupted time series

In ITS analyses, the major challenge to statistical inference is autocorrelation. Autocorrelation is when the errors in a model are correlated across time. For instance, knowing the error term in month m gives us some information about the error term in month $m + 1$. Autocorrelation results in incorrect standard errors but doesn't affect point estimates. We can test for autocorrelation by examining plots of the residuals. Usually, we can get rid of autocorrelation by accounting for seasonal factors. If this is insufficient, we can use autoregressive integrated moving average (ARIMA) models. These models predict the value of the outcome based on its own lagged value or on its previous deviation from its usual trend. Note, autocorrelation may affect DD estimators as well if data are correlated across time.

Conducting a difference-in-differences analysis in R: soda ban

Some states have tried banning soda in an attempt to reduce obesity. These data comprise two cross-sections of residents' average BMI, with one cross-section from 2010 (pre-ban) and one from 2012 (post-ban). Of the 10 states included here, 5 actually imposed a ban in 2011 (soda_ban=1), while the remaining 5 did not for reasons of free choice (soda_ban=0). Our goal is to fit a difference-in-differences model to estimate the causal impact of the soda ban.

First some housekeeping:

```
# Installing packages
if (!require("readstata13")) install.packages("readstata13")
if (!require("plyr")) install.packages("plyr")

# Loading library
library("readstata13")
library("tidyverse")
library("plyr")
library("lmtest")
library("sandwich")
library("jtools")
library("RCurl")

# read in data from foreign software
d <- read.dta13("did_data_bmi.dta")

# make a post-ban indicator
d$post <- (d$year==2012)
table(d$post)
```

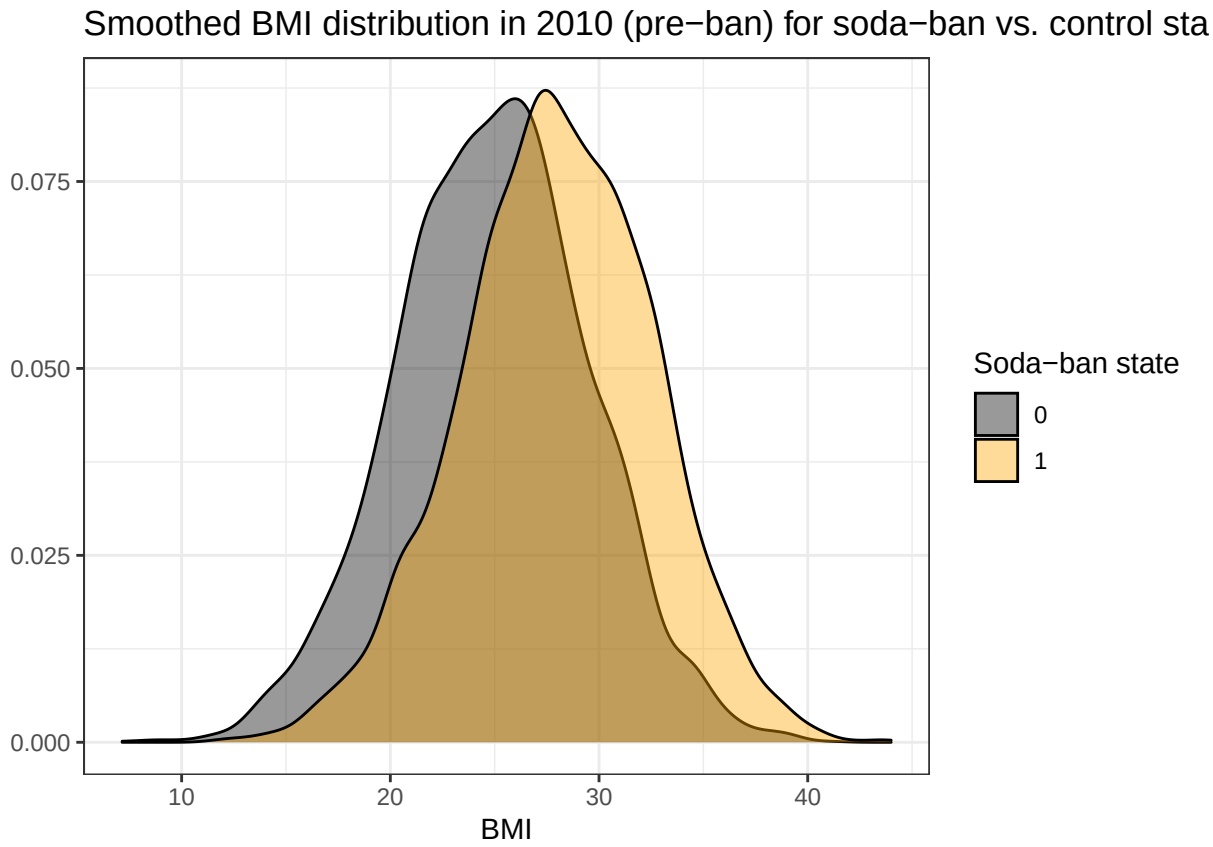
```
##
## FALSE  TRUE
## 10000  9831
```

Descriptive graphics

Let's first create a density plot of BMI in the pre-treatment period by whether or not a state eventually receives treatment (i.e., by treated and control states).

```
colors=c("black", "orange")

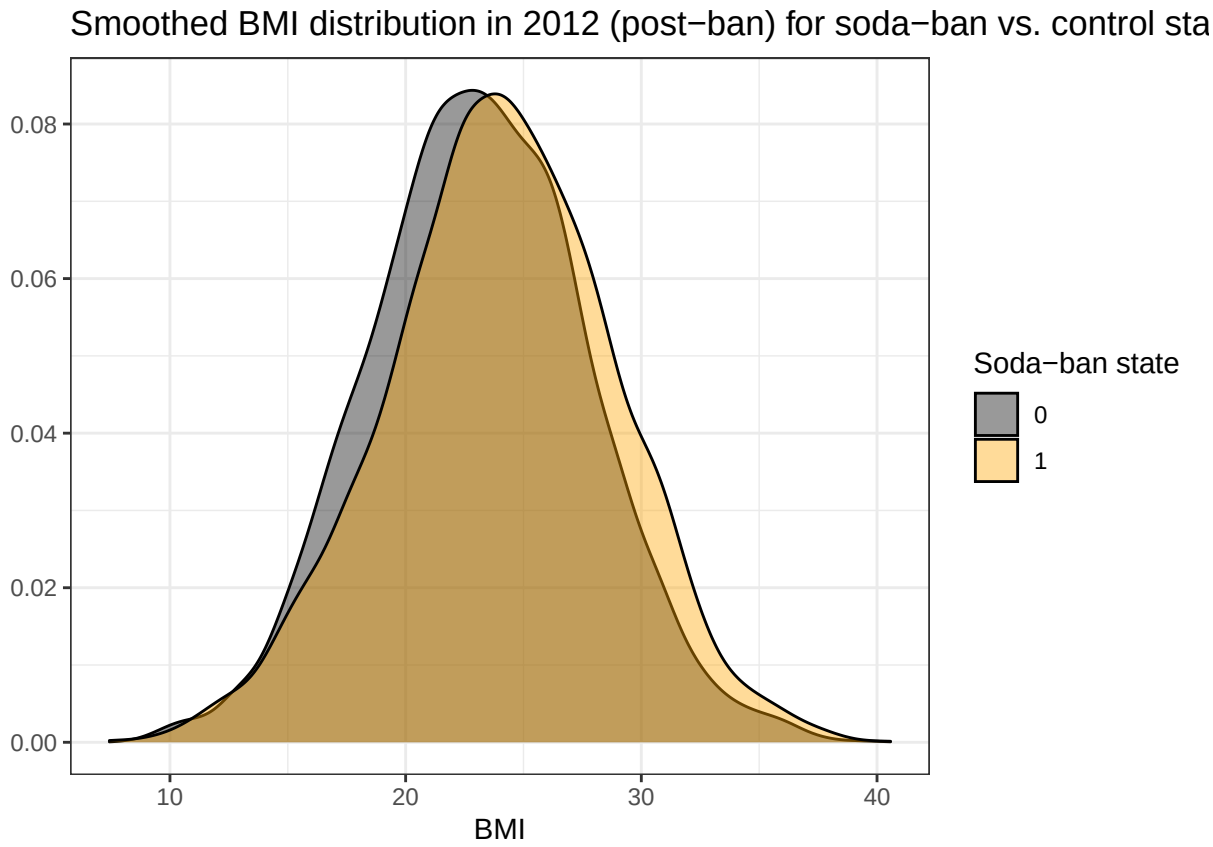
# Plotting a density plot of bmi in the pre-period
ggplot( data=d[d$post==FALSE,], aes(x=bmi, group=as.factor(soda_ban) ) ) +
  geom_density( aes( fill=as.factor(soda_ban) ), alpha=0.4 ) +
  theme_bw() + xlab("BMI") +
  ylab("") + guides(fill=guide_legend(title="Soda-ban state")) +
  scale_fill_manual(values=colors) +
  ggtitle("Smoothed BMI distribution in 2010 (pre-ban) for soda-ban vs. control states")
```



Now, let's take a look at the density of the outcome in the post-treatment period again disaggregated by treatment status.

```
colors=c("black", "orange")

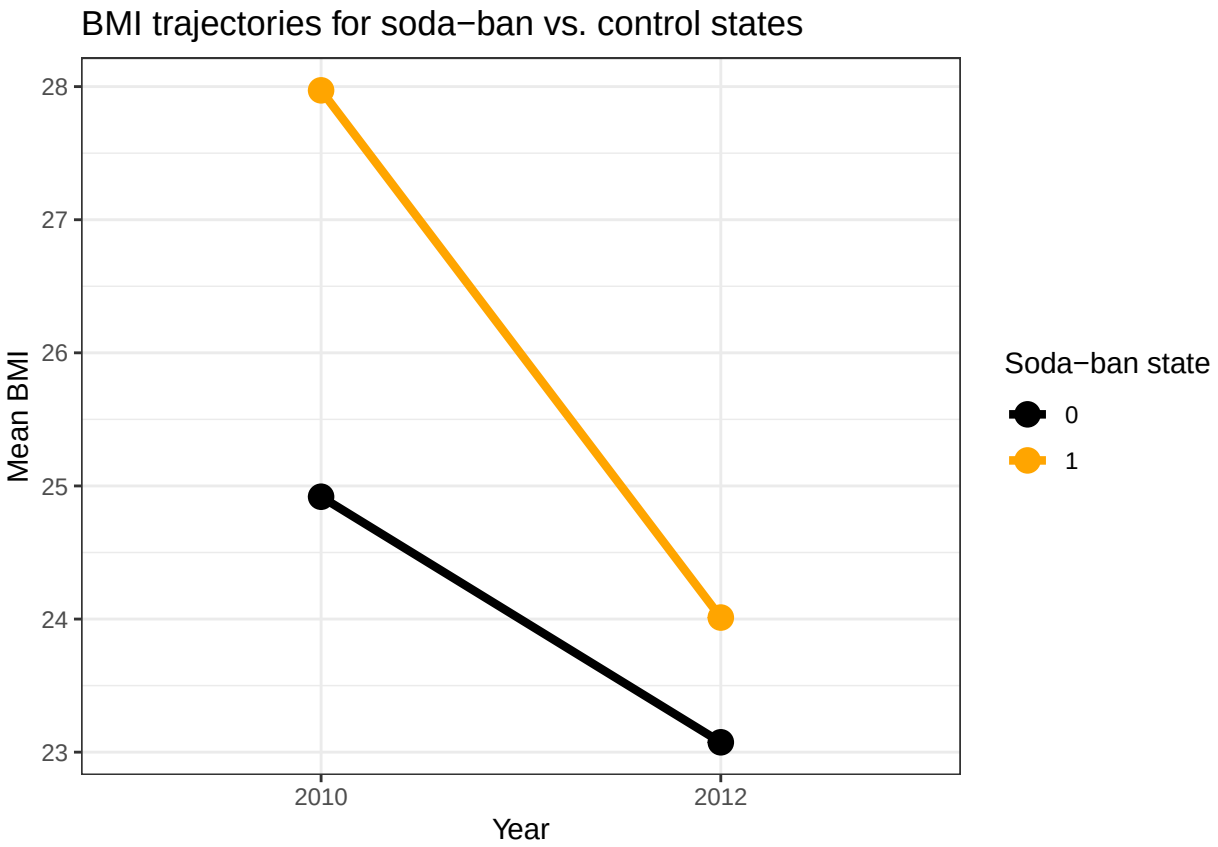
# Plotting a density plot of bmi in the pos-period
ggplot( data=d[d$post==TRUE,], aes(x=bmi, group=as.factor(soda_ban) ) ) +
  geom_density( aes( fill=as.factor(soda_ban) ), alpha=0.4 ) +
  theme_bw() + xlab("BMI") +
  ylab("") + guides(fill=guide_legend(title="Soda-ban state")) +
  scale_fill_manual(values=colors) +
  ggtitle("Smoothed BMI distribution in 2012 (post-ban) for soda-ban vs. control states")
```



Finally, let's see what the change in levels is in terms of average BMI by treated and control groups between 2010 and 2012.

```
# aggregate data for plotting average bmi
agg <- ddply(d, .(year, soda_ban), summarize, val = mean(bmi) )

# Creating a plot of the aggregated data
ggplot( data=agg, aes(x=as.factor(year), y=val, group=soda_ban,
                      color=as.factor(soda_ban) ) ) +
  geom_line(lwd=1.5) + scale_color_manual(values=colors) +
  theme_bw() + geom_point(size=4) +
  ylab("Mean BMI") + xlab("Year") +
  guides(color=guide_legend(title="Soda-ban state")) +
  ggtitle("BMI trajectories for soda-ban vs. control states")
```



Assuming the identifying assumptions hold, what might you expect the DD estimate to be from a standard DD analysis?

Basic 2x2 difference-in-differences estimation

The basic 2x2 DD model in this case will be:

$$bmi_{ist} = \alpha + \beta sodaban_s + \gamma post_t + \delta sodaban_s * post_t + u_{ist}$$

We can estimate this as follows:

```
# Estimating basic 2x2 DD model
did <- lm(bmi ~ post * soda_ban, data=d)
summ(did)
```

Observations	19831
Dependent variable	bmi
Type	OLS linear regression

F(3,19827)	1028.15
R^2	0.13
Adj. R^2	0.13

Let's interpret these coefficients:

1) Intercept: the average BMI in the states without the soda ban in 2010 is 24.92 kg/m^2.

	Est.	S.E.	t val.	p
(Intercept)	24.92	0.07	380.65	0.00
postTRUE	-1.84	0.09	-19.77	0.00
soda_ban	3.05	0.09	32.82	0.00
postTRUE:soda_ban	-2.12	0.13	-16.01	0.00

Standard errors: OLS

2) soda_ban: in 2010, on average, BMI in states that ultimately banned soda is 3.05 kg/m^2 higher than in states that did not ban soda. That is, average BMI in 2010 in states with the soda ban is $24.92 + 3.05 = 27.97$ kg/m^2.

3) post: in 2012, on average, BMI in states without the soda ban was 1.84 kg/m^2 lower than average BMI in the same states in 2010.

4) post*soda_ban: this is the **difference-in-differences estimate**. We can interpret it as follows: in 2012, BMI in states with the soda ban was, on average, an additional 2.12 kg/m^2 lower than states without the soda ban in the same year. If the DD identifying assumptions hold, then this represents the causal effect of the soda ban across the states that implemented the policy.

Correcting the standard errors

As mentioned earlier, estimating standard errors without accounting for correlated outcomes within groups is a bad idea. If we have a relatively large number of clusters, we can use the clustered standard error correction. What exactly is large in terms of number of clusters is a matter of some debate. For now, we will implement the clustered standard error correction:

```
# Installing necessary function
library(httr)
```

```
## Warning: package 'httr' was built under R version 3.6.3
```

```
url_robust <- "https://raw.githubusercontent.com/IsidoreBeautrelet/economictheoryblog/master/robust_summ
eval(parse(url_robust))

# Applying the clustered standard error correction
summary(did, cluster = c("state_id"))
```

```
##
## Call:
## lm(formula = bmi ~ post * soda_ban, data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.7720  -3.1081   0.0255   3.1554  16.6480
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       24.91895    0.67134  37.118  < 2e-16 ***
## postTRUE          -1.84489    0.02375 -77.673  < 2e-16 ***
## soda_ban           3.05309    1.04349   2.926  0.00344 **
## postTRUE:soda_ban -2.11595    0.12418 -17.039  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.652 on 19827 degrees of freedom
```

```
## Multiple R-squared:  0.1346, Adjusted R-squared:  0.1345
## F-statistic:  5004 on 3 and 9 DF,  p-value: 8.156e-15
```

Estimating the difference-in-differences model with group fixed effects

Suppose we wanted to account for time-invariant, state-specific factors because we believe this will allow us to better account for differences in average BMI across states at baseline. We can easily estimate this model using state fixed effects!

```
# Creating an interaction term
d$post.soda_ban <- d$post * d$soda_ban

# Applying state fixed effects using dummy coding
did_fe <- lm(bmi ~ as.factor(state_id) + post + post.soda_ban, data=d)
summ(did_fe)
```

Observations	19831
Dependent variable	bmi
Type	OLS linear regression

F(11,19819)	571.23
R^2	0.24
Adj. R^2	0.24

	Est.	S.E.	t val.	p
(Intercept)	22.96	0.11	215.88	0.00
as.factor(state_id)2	0.77	0.14	5.59	0.00
as.factor(state_id)3	1.58	0.15	10.36	0.00
as.factor(state_id)4	2.20	0.14	15.90	0.00
as.factor(state_id)5	2.92	0.14	21.13	0.00
as.factor(state_id)6	3.88	0.14	28.13	0.00
as.factor(state_id)7	4.76	0.15	31.58	0.00
as.factor(state_id)8	5.37	0.15	35.52	0.00
as.factor(state_id)9	6.36	0.15	41.84	0.00
as.factor(state_id)10	6.76	0.15	45.27	0.00
postTRUE	-1.84	0.09	-21.03	0.00
post.soda_ban	-2.08	0.12	-16.81	0.00

Standard errors: OLS

```
# We can apply an additional clustering correction to the SEs as well
summary(did_fe, cluster = "state_id")
```

```
##
## Call:
## lm(formula = bmi ~ as.factor(state_id) + post + post.soda_ban,
##     data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.8097  -2.9721   0.0182   2.9417  18.1646
##
```

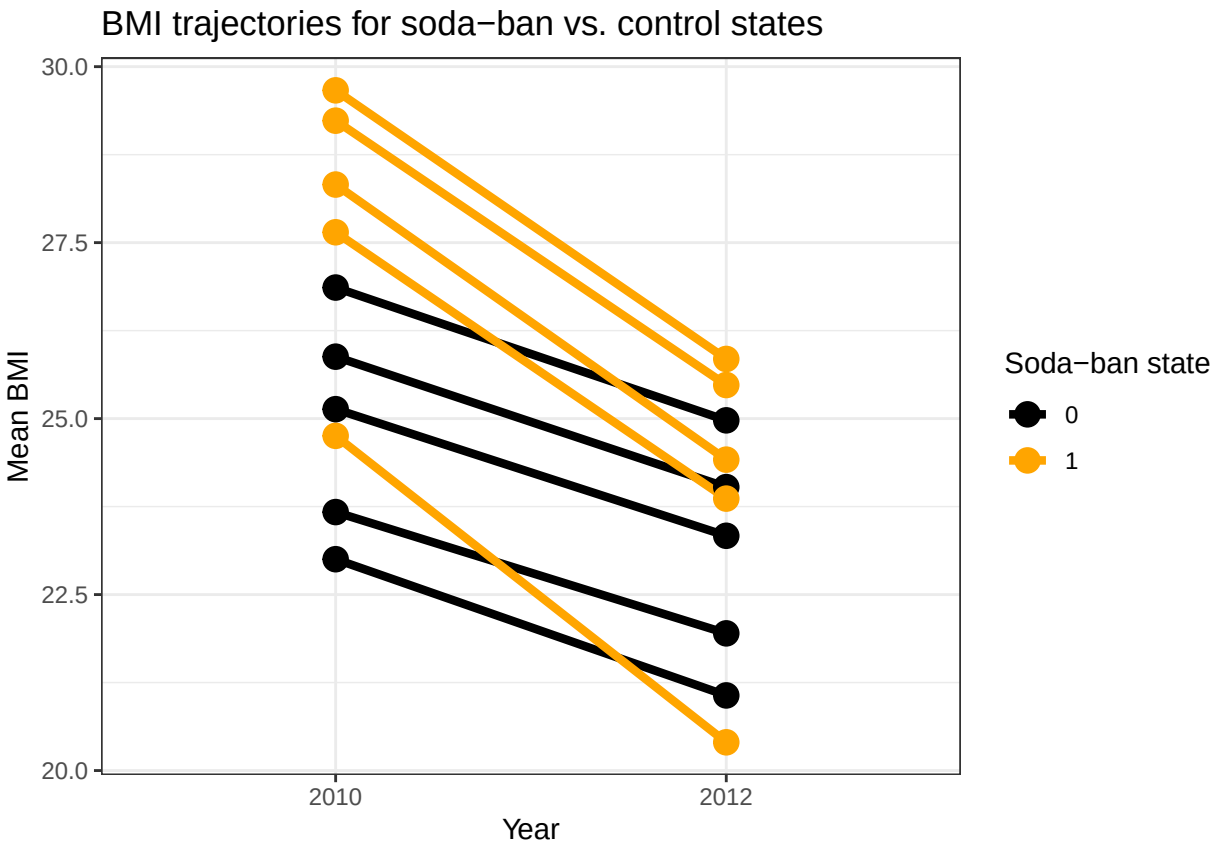
```
## Coefficients:
##                          Estimate Std. Error  t value Pr(>|t|)
## (Intercept)           22.9565950  0.0164745  1393.46   <2e-16 ***
## as.factor(state_id)2   0.7747327  0.0009701   798.65   <2e-16 ***
## as.factor(state_id)3   1.5781667  0.0544514    28.98   <2e-16 ***
## as.factor(state_id)4   2.1980834  0.0008980  2447.80   <2e-16 ***
## as.factor(state_id)5   2.9169217  0.0001520 19190.81   <2e-16 ***
## as.factor(state_id)6   3.8821253  0.0004331  8964.37   <2e-16 ***
## as.factor(state_id)7   4.7604698  0.0545305    87.30   <2e-16 ***
## as.factor(state_id)8   5.3746748  0.0526709   102.04   <2e-16 ***
## as.factor(state_id)9   6.3558515  0.0520136   122.20   <2e-16 ***
## as.factor(state_id)10  6.7596433  0.0539149   125.38   <2e-16 ***
## postTRUE              -1.8394846  0.0343478   -53.55   <2e-16 ***
## post.soda_ban         -2.0811078  0.1076209   -19.34   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.358 on 19819 degrees of freedom
## Multiple R-squared:  0.2407, Adjusted R-squared:  0.2403
## F-statistic: 3.268e+23 on 11 and 9 DF,  p-value: < 2.2e-16
```

Note that we did not include a "main effect" for state in our DD model. The reason for this is that a state "main effect" would be perfectly collinear with the state fixed effects. We could have included state in the model, but it would have just been dropped from the estimation due to perfect collinearity! Including or not including the state variable does not affect our DD estimate.

Finally, if we want to visualize what our state FE model looks like:

```
# Creating a plot of the state FE model
agg2 = ddply(d, .(year, state_id, soda_ban), summarize, val = mean(bmi) )

ggplot( data=agg2, aes(x=as.factor(year), y=val, group=state_id,
                       color=as.factor(soda_ban) ) ) +
  geom_line(lwd=1.5) + scale_color_manual(values=colors) +
  theme_bw() + geom_point(size=4) +
  ylab("Mean BMI") + xlab("Year") +
  guides(color=guide_legend(title="Soda-ban state")) +
  ggtitle("BMI trajectories for soda-ban vs. control states")
```


BMI trajectories for soda−ban vs. control states
Mean BMI
30.0
27.5
25.0
22.5
20.0
2010
2012
Year
Soda−ban state
0
1

Conducting a simple ITS analysis in R: Italy's smoking ban

The code below is from the tutorial presented by Bernal et. al. (2017). The data itself come from a study by Barone-Adesi et. al. (2011). The treatment of interest is a smoking ban in public places introduced nation-wide in Italy on January 10, 2005. The outcome of interest is Acute Coronary Events (ACE), which are coded under ICD10 codes 410 (acute myocardial infarctions) and 411 (acute and subacute forms of ischemic heart disease). The authors of the original study argue that ACE is a short term outcome with rapid onset, and is therefore suitable for an ITS analysis. The tutorial presented in Bernal et. al. (2017) uses data from the Sciily region between 2002 and 2006 (59 months) and includes monthly counts of ACEs for all individuals between the ages of 0-69 years.

Let's start by loading the relevant packages and the data.

```
# Installing packages
if (!require("foreign")) install.packages("foreign")
if (!require("tsModel")) install.packages("tsModel")
if (!require("lmtest")) install.packages("lmtest")
if (!require("Epi")) install.packages("Epi")
if (!require("splines")) install.packages("splines")
if (!require("vcd")) install.packages("vcd")

# Loading libraries
library("foreign")
library("tsModel")
library("lmtest")
library("Epi")
library("splines")
library("vcd")

# Loading data
data <- read.csv("sicily.csv")

# Looking at the first few observations
head(data)
```

```
##   year month aces time smokban      pop   stdpop
## 1 2002     1  728    1       0 364277.4 379875.3
## 2 2002     2  659    2       0 364277.4 376495.5
## 3 2002     3  791    3       0 364277.4 377040.8
## 4 2002     4  734    4       0 364277.4 377116.4
## 5 2002     5  757    5       0 364277.4 377383.4
## 6 2002     6  726    6       0 364277.4 374113.1
```

This dataset includes the following variables:

- year
- month
- time = elapsed time since the start of the study
- aces = count of acute coronary episodes 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 (in person-years)

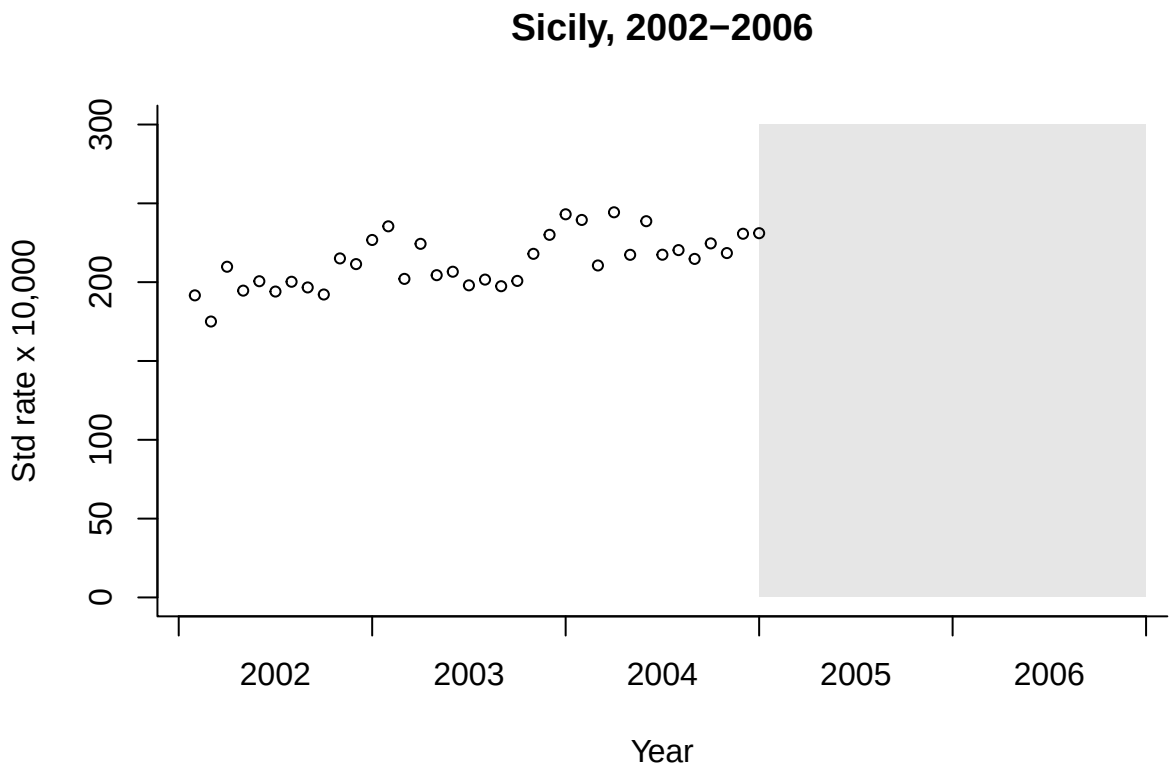
Descriptive analyses

As a first step, we'll examine the data. We will look at the pre-trend to get an indication of how stable the trend is over time, whether a linear model is likely to be appropriate, and whether there appears to be

seasonality.

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

# start the plot, excluding the points and the x-axis
{plot(data$rate,type="n",ylim=c(00,300),xlab="Year", ylab="Std rate x 10,000",
bty="l",xaxt="n")
# shade the post intervention period grey
rect(36,0,60,300,col=grey(0.9),border=F)
# plot the observed rate for pre-intervention period
points(data$rate[data$smokban==0],cex=0.7)
#specify the x-axis (i.e. time units)
axis(1,at=0:5*12,labels=F)
axis(1,at=0:4*12+6,tick=F,labels=2002:2006)
# add a title
title("Sicily, 2002-2006")}
```




```
# Computing summary statistics
summary(data)
```

```
##       year          month            aces            time
##  Min.   :2002   Min.   : 1.000   Min.   : 659.0   Min.   : 1.0
##  1st Qu.:2003   1st Qu.: 3.500   1st Qu.: 766.5   1st Qu.:15.5
##  Median :2004   Median : 6.000   Median : 831.0   Median :30.0
##  Mean   :2004   Mean   : 6.407   Mean   : 829.1   Mean   :30.0
##  3rd Qu.:2005   3rd Qu.: 9.000   3rd Qu.: 878.0   3rd Qu.:44.5
```

```
##  Max.   :2006   Max.   :12.000   Max.   :1021.0   Max.   :59.0
##     smokban            pop             stdpop           rate
##  Min.   :0.0000   Min.   :363351   Min.   :374113   Min.   :175.0
##  1st Qu.:0.0000   1st Qu.:363833   1st Qu.:379437   1st Qu.:201.4
##  Median :0.0000   Median :364277   Median :383428   Median :215.4
##  Mean   :0.3898   Mean   :364121   Mean   :383464   Mean   :216.1
##  3rd Qu.:1.0000   3rd Qu.:364421   3rd Qu.:387335   3rd Qu.:227.7
##  Max.   :1.0000   Max.   :364700   Max.   :394006   Max.   :260.7
```

```
# Tabulating ACES before and after the smoking ban
summary(data$aces[data$smokban==0])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   659.0   759.2   811.0   810.9   858.5   937.0
```

```
summary(data$aces[data$smokban==1])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   738.0   815.0   843.0   857.5   898.0  1021.0
```

```
summary(data$rate[data$smokban==0])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   175.0   200.5   213.1   213.3   225.1   244.3
```

```
summary(data$rate[data$smokban==1])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   190.9   209.6   217.9   220.5   231.0   260.7
```

Poisson regression model

Next we are going to model the outcome using an interrupted time-series specification with a change in intercept (but no change in slope). We will use a Poisson model to model the ACE count directly. We will use the standardized population (log transformed) to transform back to rates.

```
model1 <- glm(aces ~ offset(log(stdpop)) + smokban + time, family=poisson, data)
summary(model1)
```

```
##
## Call:
## glm(formula = aces ~ offset(log(stdpop)) + smokban + time, family = poisson,
##     data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.4336  -1.4432  -0.4431   1.2735   5.0701
##
## Coefficients:
##               Estimate Std. Error  z value Pr(>|z|)
## (Intercept) -6.2431509  0.0111894 -557.955  < 2e-16 ***
## smokban     -0.1116284  0.0172248   -6.481 9.13e-11 ***
## time         0.0049450  0.0004992    9.905  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
```

```
##     Null deviance: 308.52  on 58  degrees of freedom
## Residual deviance: 197.31  on 56  degrees of freedom
## AIC: 708.03
##
## Number of Fisher Scoring iterations: 3
```

```
summary(model1)$dispersion # note that dispersion is forced to 1 here (variance = mean)
```

```
## [1] 1
```

```
round(ci.lin(model1,Exp=T),3) # compute confidence intervals
```

```
##             Estimate StdErr        z P exp(Est.)  2.5% 97.5%
## (Intercept)   -6.243  0.011 -557.955 0     0.002 0.002 0.002
## smokban       -0.112  0.017   -6.481 0     0.894 0.865 0.925
## time           0.005  0.000    9.905 0     1.005 1.004 1.006
```

```
# Create a new dataframe with more time units to improve the graph
datanew <- data.frame(stdpop=mean(data$stdpop),smokban=rep(c(0,1),c(360,240)),
  time= 1:600/10,month=rep(1:120/10,5))

# Generate predicted values based on the model in order to create a plot
pred1 <- predict(model1,type="response",datanew)/mean(data$stdpop)*10^5

# Plot predictions on the scatter graph
{plot(data$rate,type="n",ylim=c(0,300),xlab="Year",ylab="Std rate x 10,000",
  bty="l",xaxt="n")
rect(36,0,60,300,col=grey(0.9),border=F)
points(data$rate,cex=0.7)
axis(1,at=0:5*12,labels=F)
axis(1,at=0:4*12+6,tick=F,labels=2002:2006)
lines((1:600/10),pred1,col=2)
title("Sicily, 2002-2006")

# To plot the counterfactual scenario,  create a data frame as if smokban
# (the intervention) was never being implemented
datanew <- data.frame(stdpop=mean(data$stdpop),smokban=0,time=1:600/10,
  month=rep(1:120/10,5))

# Generate predictions under the counterfactual scenario and add them to the plot
pred1b <- predict(model1,datanew,type="response")/mean(data$stdpop)*10^5
lines(datanew$time,pred1b,col=2,lty=2)}
```




```
# Return the data frame to the scenario including the intervention
datanew <- data.frame(stdpop=mean(data$stdpop),smokban=rep(c(0,1),c(360,240)),
  time= 1:600/10,month=rep(1:120/10,5))
```

Methodological issues

```
#a) Overdispersion: Quasi-Poisson model

# In the model above we have not allowed for overdispersion - in order
# to do this we can use a quasipoisson model, which allows the variance to
# be proportional rather than equal to the mean.

model2 <- glm(aces ~ offset(log(stdpop)) + smokban + time, family=quasipoisson,
  data)
summary(model2)
```

```
##
## Call:
## glm(formula = aces ~ offset(log(stdpop)) + smokban + time, family = quasipoisson,
##     data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.4336  -1.4432  -0.4431   1.2735   5.0701
##
## Coefficients:
```

```
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept) -6.2431509  0.0211096 -295.749  < 2e-16 ***
## smokban     -0.1116284  0.0324959   -3.435  0.00112 **
## time         0.0049450  0.0009418    5.250 2.43e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 3.559195)
##
##     Null deviance: 308.52  on 58  degrees of freedom
## Residual deviance: 197.31  on 56  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 3
```

```
summary(model2)$dispersion # now the variance is 3.6x the mean
```

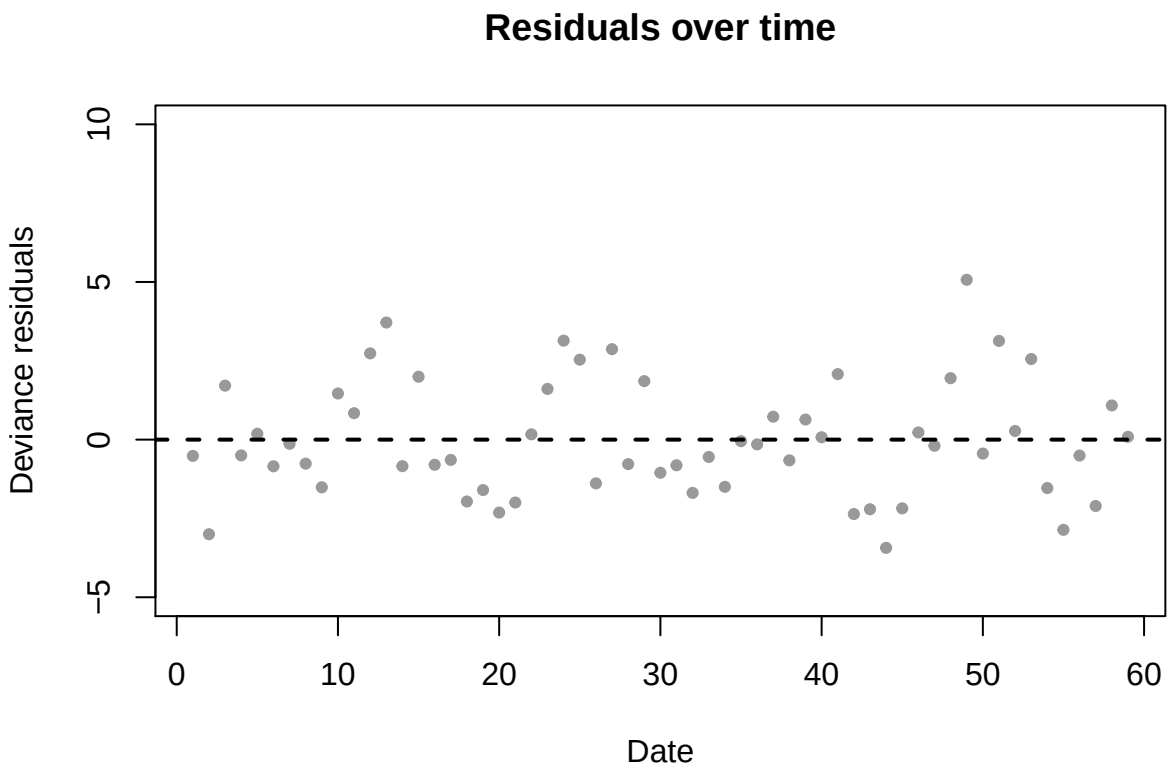
```
## [1] 3.559195
```

```
round(ci.lin(model2,Exp=T),3)
```

```
##             Estimate StdErr        z     P exp(Est.)  2.5% 97.5%
## (Intercept)   -6.243  0.021 -295.749 0.000     0.002 0.002 0.002
## smokban       -0.112  0.032   -3.435 0.001     0.894 0.839 0.953
## time           0.005  0.001    5.250 0.000     1.005 1.003 1.007
```

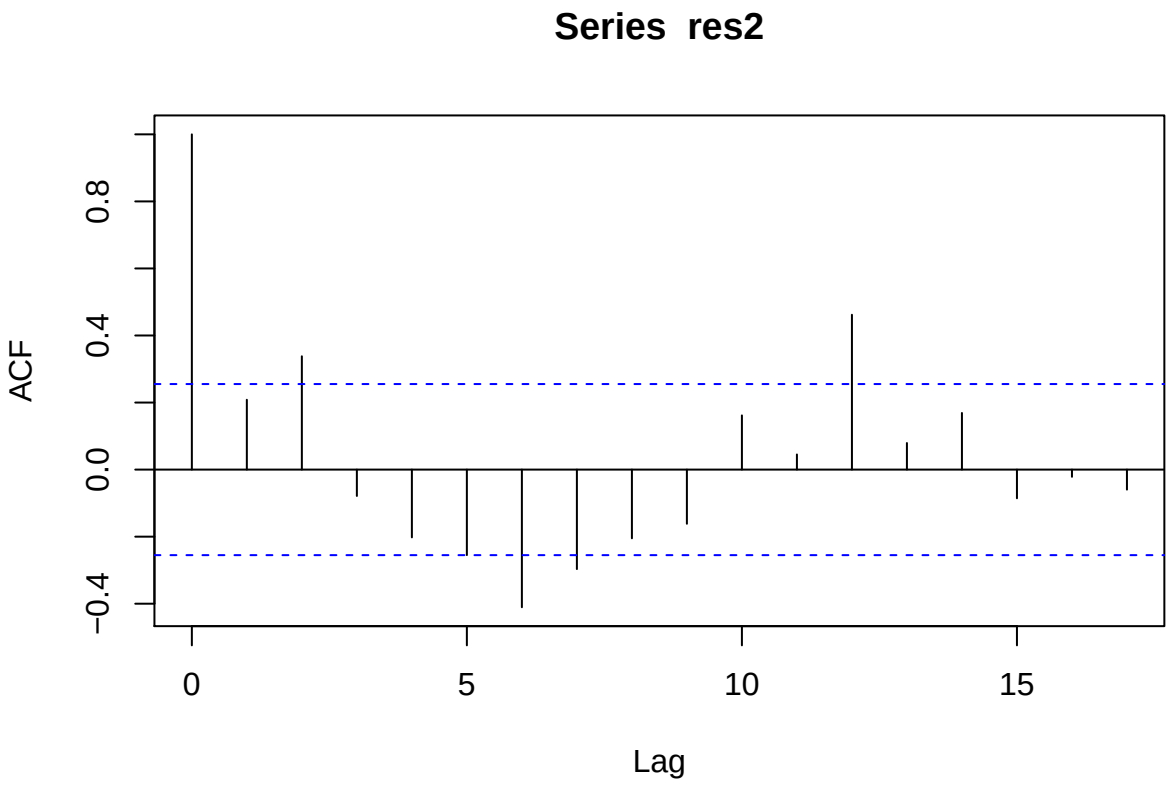
```
#b) Model checking and autocorrelation

# Check the residuals by plotting against time
res2 <- residuals(model2,type="deviance")
plot(data$time,res2,ylim=c(-5,10),pch=19,cex=0.7,col=grey(0.6),
  main="Residuals over time",ylab="Deviance residuals",xlab="Date")
abline(h=0,lty=2,lwd=2)
```

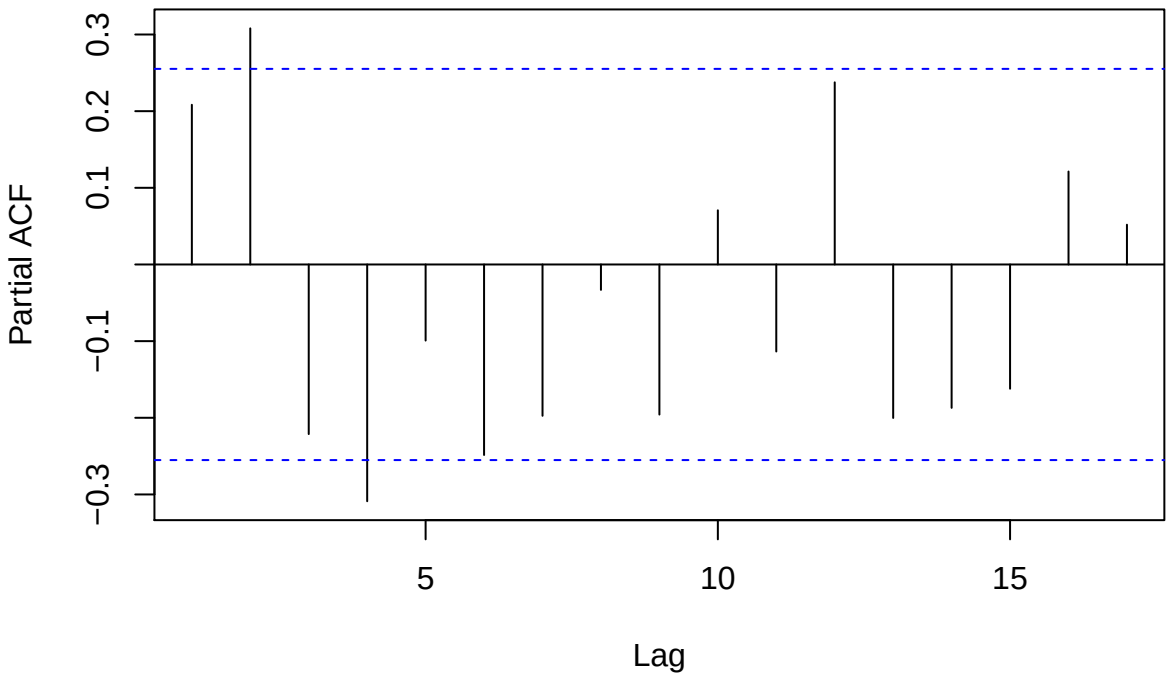



```
# Further check for autocorrelation by examining the autocorrelation and
# partial autocorrelation functions. We're looking for autocorrelations
# outside the 5% significant bands (marked by blue dotted lines)

# acf shows the coefficients of correlation between the time
# series and lags of itself.
acf(res2)
```




```
# pacf shows the partial correlation coefficients between the
# time series and lags of itself (how much correlation between Y_t and
# Y_{t-1} is not explained by their mutual correlation with other variables).
pacf(res2)
```




```
# There appears to be autocorrelation. What happens if we add in calendar-month
# dummy variables?

model2b <- glm(aces ~ offset(log(stdpop)) + smokban + time + factor(month), family=quasipoisson,
  data)
summary(model2b)
```

```
##
## Call:
## glm(formula = aces ~ offset(log(stdpop)) + smokban + time + factor(month),
##     family = quasipoisson, data = data)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.51295  -0.56864  -0.01395   0.67146   2.42379
##
## Coefficients:
##                  Estimate Std. Error  t value Pr(>|t|)
## (Intercept)     -6.171091   0.020135 -306.490  < 2e-16 ***
## smokban         -0.125008   0.021254   -5.882 4.68e-07 ***
## time             0.005506   0.000625    8.810 2.36e-11 ***
## factor(month)2  -0.125726   0.025756   -4.881 1.36e-05 ***
## factor(month)3  -0.010847   0.024991   -0.434 0.666353
## factor(month)4  -0.094599   0.025544   -3.703 0.000579 ***
## factor(month)5  -0.040625   0.025223   -1.611 0.114256
## factor(month)6  -0.138879   0.025908   -5.360 2.75e-06 ***
```

```
## factor(month)7  -0.138513   0.025919   -5.344 2.90e-06 ***
## factor(month)8  -0.146162   0.026057   -5.609 1.18e-06 ***
## factor(month)9  -0.144057   0.026097   -5.520 1.60e-06 ***
## factor(month)10 -0.076388   0.025767   -2.965 0.004834 **
## factor(month)11 -0.071298   0.025825   -2.761 0.008313 **
## factor(month)12 -0.021458   0.027031   -0.794 0.431463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 1.365659)
##
##     Null deviance: 308.517  on 58  degrees of freedom
## Residual deviance:  61.732  on 45  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 3
```

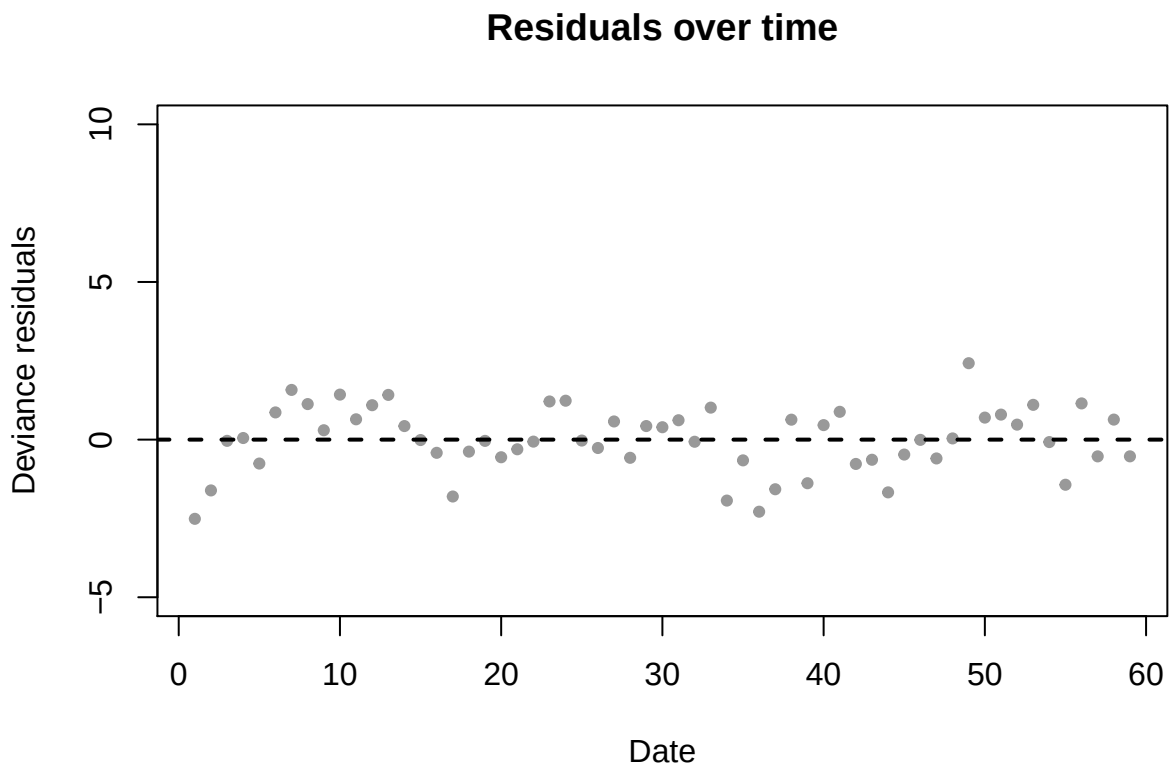
```
summary(model2b)$dispersion # now the variance is 1.4x the mean
```

```
## [1] 1.365659
```

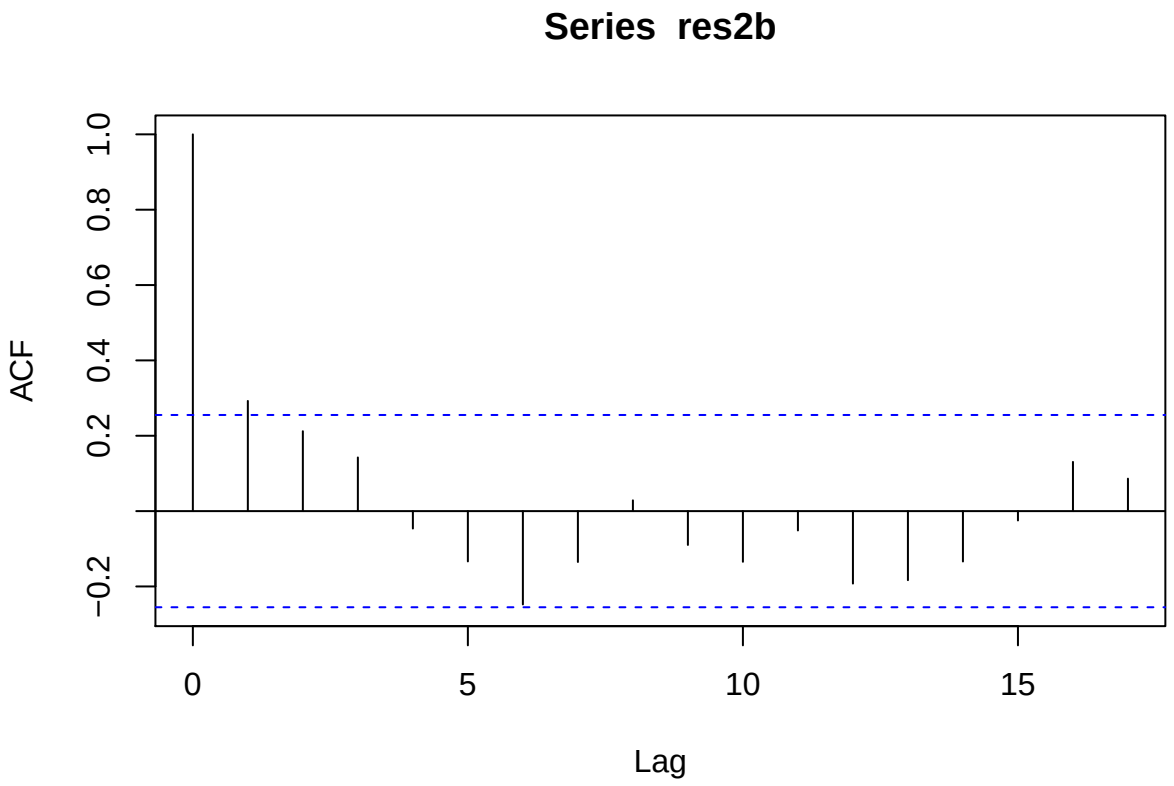
```
round(ci.lin(model2b,Exp=T),3)
```

```
##                 Estimate StdErr        z     P exp(Est.)  2.5% 97.5%
## (Intercept)       -6.171  0.020 -306.490 0.000     0.002 0.002 0.002
## smokban           -0.125  0.021   -5.882 0.000     0.882 0.846 0.920
## time               0.006  0.001    8.810 0.000     1.006 1.004 1.007
## factor(month)2    -0.126  0.026   -4.881 0.000     0.882 0.838 0.928
## factor(month)3    -0.011  0.025   -0.434 0.664     0.989 0.942 1.039
## factor(month)4    -0.095  0.026   -3.703 0.000     0.910 0.865 0.956
## factor(month)5    -0.041  0.025   -1.611 0.107     0.960 0.914 1.009
## factor(month)6    -0.139  0.026   -5.360 0.000     0.870 0.827 0.916
## factor(month)7    -0.139  0.026   -5.344 0.000     0.871 0.828 0.916
## factor(month)8    -0.146  0.026   -5.609 0.000     0.864 0.821 0.909
## factor(month)9    -0.144  0.026   -5.520 0.000     0.866 0.823 0.911
## factor(month)10   -0.076  0.026   -2.965 0.003     0.926 0.881 0.974
## factor(month)11   -0.071  0.026   -2.761 0.006     0.931 0.885 0.980
## factor(month)12   -0.021  0.027   -0.794 0.427     0.979 0.928 1.032
```

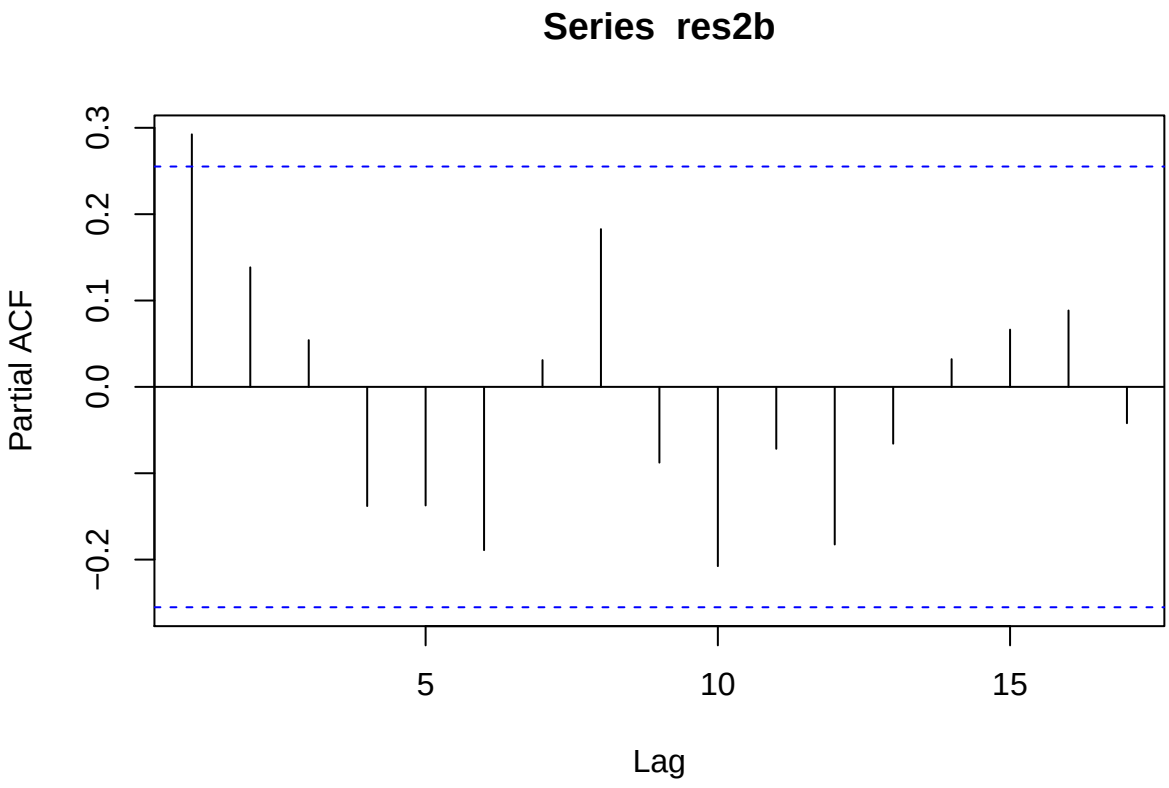
```
res2b <- residuals(model2b,type="deviance")
plot(data$time,res2b,ylim=c(-5,10),pch=19,cex=0.7,col=grey(0.6),
  main="Residuals over time",ylab="Deviance residuals",xlab="Date")
abline(h=0,lty=2,lwd=2)
```


Residuals over time
Deviance residuals
10
5
0
−5
0
10
20
30
40
50
60
Date

```
acf(res2b)
```




```
pacf(res2b)
```




```
# This looks better!
```

Adjusting for seasonality

```
# There are various ways of adjusting for seasonality.
# Here we use harmonic terms specifying the number of sin and cosine
# pairs to include (in this case 2) and the length of the period
# (12 months). This type of model allows for multiple seasonality.
model3 <- glm(aces ~ offset(log(stdpop)) + smokban + time +
  harmonic(month,2,12), family=quasipoisson, data)
summary(model3)
```

```
##
## Call:
## glm(formula = aces ~ offset(log(stdpop)) + smokban + time + harmonic(month,
##     2, 12), family = quasipoisson, data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.5292  -1.2634  -0.0972   1.2571   3.6882
##
## Coefficients:
##                           Estimate Std. Error  t value Pr(>|t|)
## (Intercept)            -6.2523723  0.0175689 -355.877  < 2e-16 ***
## smokban                -0.1220868  0.0270311   -4.517 3.64e-05 ***
## time                    0.0053892  0.0007955    6.774 1.13e-08 ***
```

```
## harmonic(month, 2, 12)1  0.0370967  0.0100241    3.701 0.000520 ***
## harmonic(month, 2, 12)2 -0.0182122  0.0096435   -1.889 0.064537 .
## harmonic(month, 2, 12)3  0.0381608  0.0096848    3.940 0.000244 ***
## harmonic(month, 2, 12)4  0.0147634  0.0097221    1.519 0.134935
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 2.262186)
##
##     Null deviance: 308.52  on 58  degrees of freedom
## Residual deviance: 117.27  on 52  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 3
```

```
summary(model3)$dispersion
```

```
## [1] 2.262186
```

```
round(ci.lin(model3,Exp=T),3)
```

```
##                         Estimate StdErr        z     P exp(Est.)  2.5% 97.5%
## (Intercept)               -6.252  0.018 -355.877 0.000     0.002 0.002 0.002
## smokban                   -0.122  0.027   -4.517 0.000     0.885 0.839 0.933
## time                       0.005  0.001    6.774 0.000     1.005 1.004 1.007
## harmonic(month, 2, 12)1    0.037  0.010    3.701 0.000     1.038 1.018 1.058
## harmonic(month, 2, 12)2   -0.018  0.010   -1.889 0.059     0.982 0.964 1.001
## harmonic(month, 2, 12)3    0.038  0.010    3.940 0.000     1.039 1.019 1.059
## harmonic(month, 2, 12)4    0.015  0.010    1.519 0.129     1.015 0.996 1.034
```

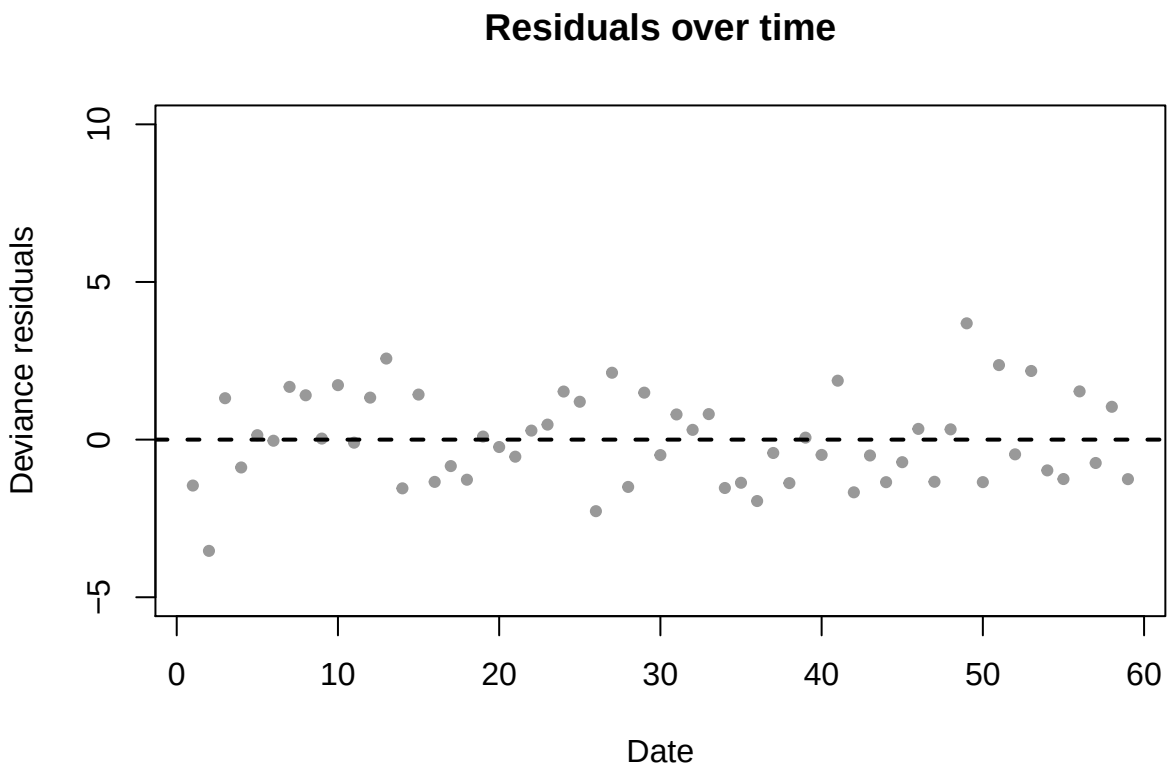
```
# EFFECTS
ci.lin(model3,Exp=T)["smokban",5:7]
```

```
## exp(Est.)      2.5%     97.5%
## 0.8850715 0.8394010 0.9332270
```

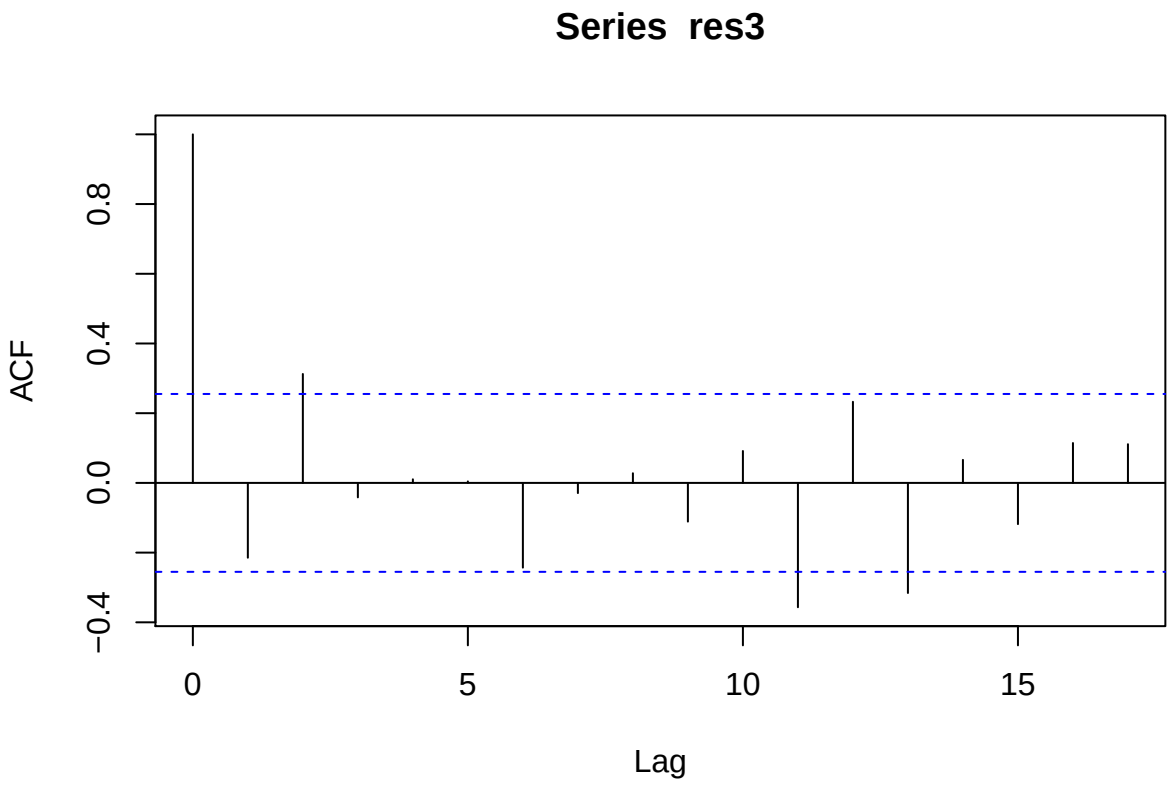
```
# TREND
exp(coef(model3)["time"]*12)
```

```
##     time
## 1.066808
```

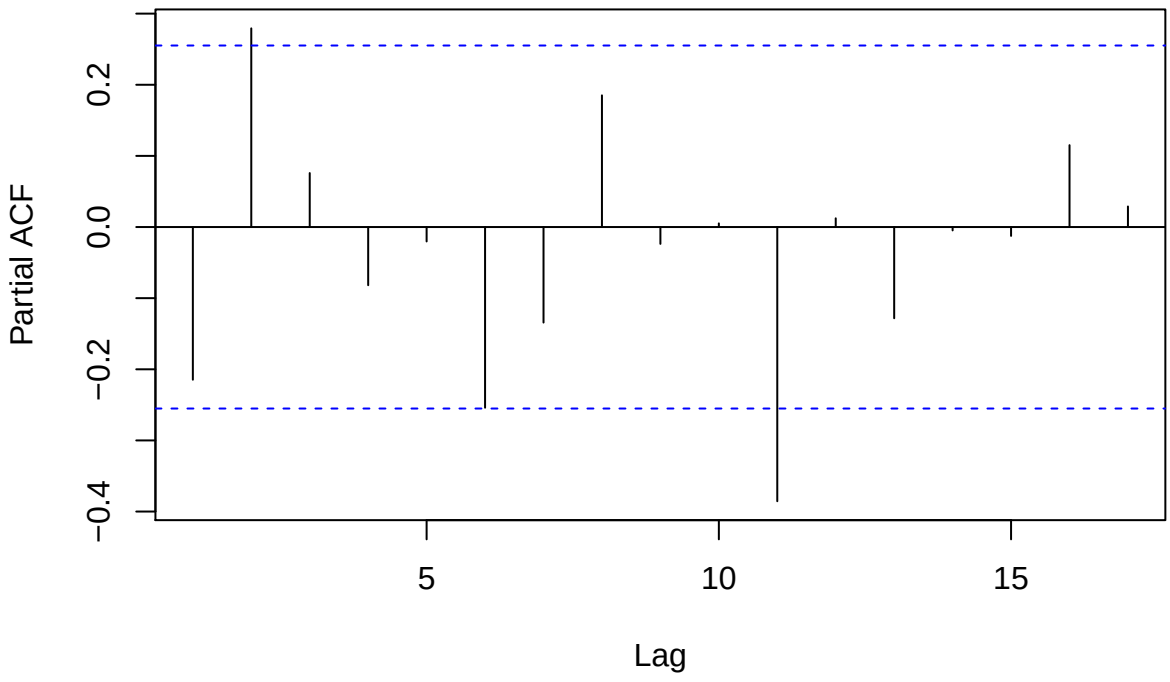
```
# We again check the model and autocorrelation functions
res3 <- residuals(model3,type="deviance")
plot(res3,ylim=c(-5,10),pch=19,cex=0.7,col=grey(0.6),main="Residuals over time",
  ylab="Deviance residuals",xlab="Date")
abline(h=0,lty=2,lwd=2)
```


Residuals over time
Deviance residuals
10
5
0
−5
0
10
20
30
40
50
60
Date

```
acf(res3)
```



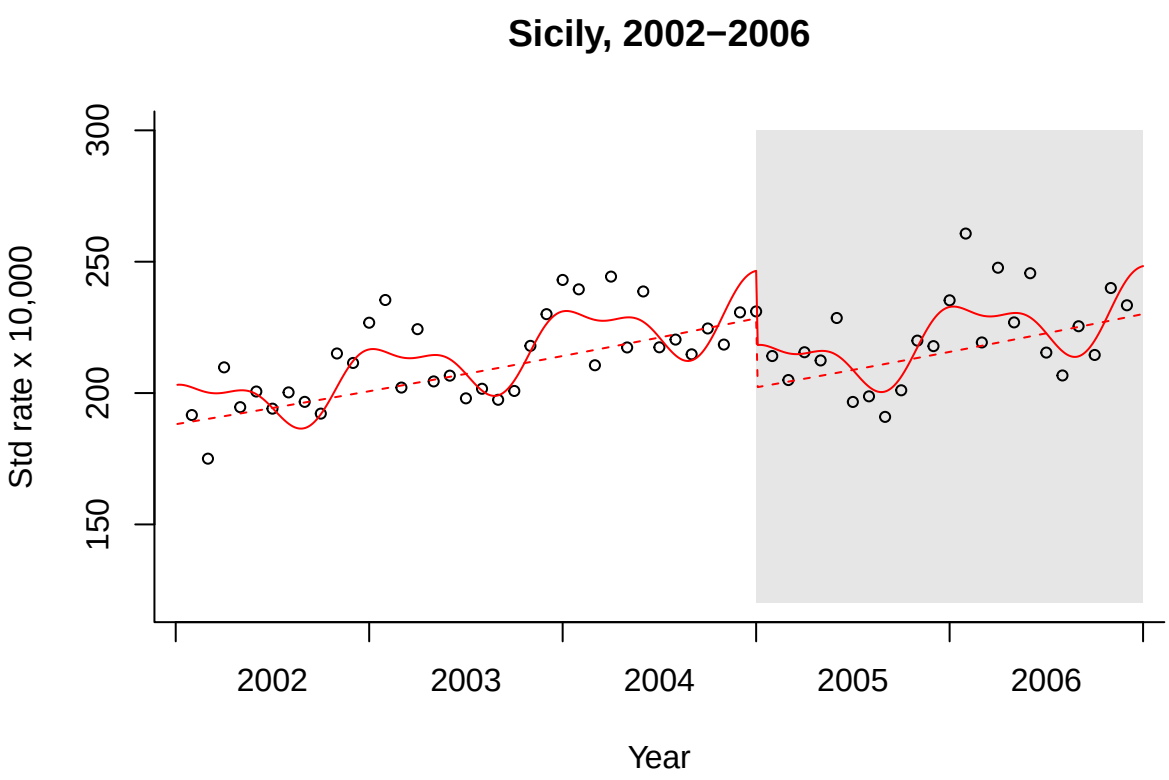

```
pacf(res3)
```




```
# Predict and plot of the seasonally adjusted model
pred3 <- predict(model3,type="response",datanew)/mean(data$stdpop)*10^5
{plot(data$rate,type="n",ylim=c(120,300),xlab="Year",ylab="Std rate x 10,000",
  bty="l",xaxt="n")
rect(36,120,60,300,col=grey(0.9),border=F)
points(data$rate,cex=0.7)
axis(1,at=0:5*12,labels=F)
axis(1,at=0:4*12+6,tick=F,labels=2002:2006)
lines(1:600/10,pred3,col=2)
title("Sicily, 2002-2006")

# It is sometimes difficult to clearly see the change graphically in the
# seasonally adjusted model, therefore it can be useful to plot a straight
# line representing a 'deseasonalised' trend. This can be done by
# predicting all the observations for the same month. In this case, we use June.
pred3b <- predict(model3,type="response",transform(datanew,month=6))/
  mean(data$stdpop)*10^5

# This can then be added to the plot as a dashed line.
lines(1:600/10,pred3b,col=2,lty=2)}
```



Adding a change-in-slope

```
# We parameterize the change-in-slope as an interaction between time
# and the ban indicator.
model4 <- glm(aces ~ offset(log(stdpop)) + smokban*time + harmonic(month,2,12),
  family=quasipoisson, data)
summary(model4)
```

```
##
## Call:
## glm(formula = aces ~ offset(log(stdpop)) + smokban * time + harmonic(month,
##     2, 12), family = quasipoisson, data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.7028  -1.1977  -0.1371   1.1837   3.6172
##
## Coefficients:
##                            Estimate Std. Error  t value Pr(>|t|)
## (Intercept)              -6.2469039  0.0188768 -330.930  < 2e-16 ***
## smokban                  -0.1852194  0.0835610   -2.217 0.031138 *
## time                      0.0051026  0.0008745    5.835 3.72e-07 ***
## harmonic(month, 2, 12)1   0.0383261  0.0101687    3.769 0.000427 ***
## harmonic(month, 2, 12)2 -0.0176326  0.0096966   -1.818 0.074872 .
## harmonic(month, 2, 12)3   0.0383558  0.0097142    3.948 0.000242 ***
## harmonic(month, 2, 12)4   0.0149653  0.0097524    1.535 0.131081
```

```
## smokban:time             0.0014822  0.0018546    0.799 0.427879
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 2.274563)
##
##     Null deviance: 308.52  on 58  degrees of freedom
## Residual deviance: 115.81  on 51  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 3
```

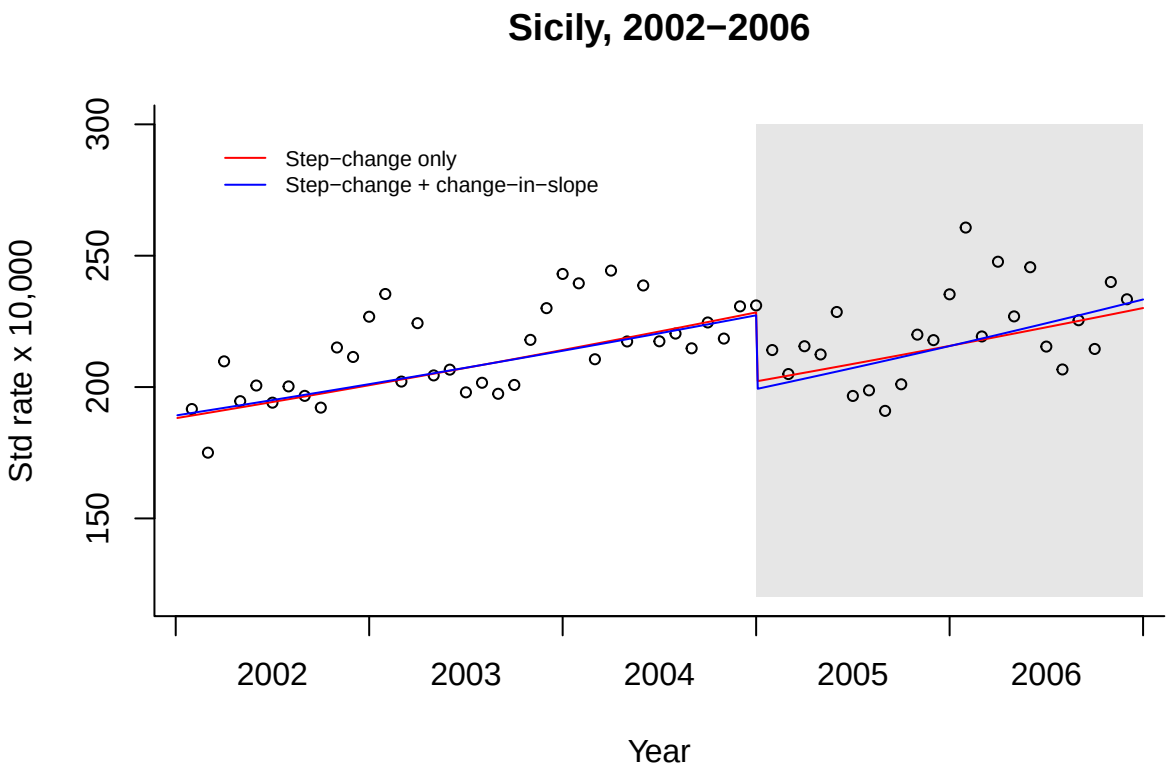
```
round(ci.lin(model4,Exp=T),3)
```

```
##                          Estimate StdErr        z     P exp(Est.)  2.5% 97.5%
## (Intercept)                -6.247  0.019 -330.930 0.000     0.002 0.002 0.002
## smokban                    -0.185  0.084   -2.217 0.027     0.831 0.705 0.979
## time                        0.005  0.001    5.835 0.000     1.005 1.003 1.007
## harmonic(month, 2, 12)1     0.038  0.010    3.769 0.000     1.039 1.019 1.060
## harmonic(month, 2, 12)2    -0.018  0.010   -1.818 0.069     0.983 0.964 1.001
## harmonic(month, 2, 12)3     0.038  0.010    3.948 0.000     1.039 1.020 1.059
## harmonic(month, 2, 12)4     0.015  0.010    1.535 0.125     1.015 0.996 1.035
## smokban:time                0.001  0.002    0.799 0.424     1.001 0.998 1.005
```

```
# Predict and plot the 'deseasonalised' trend

# Compare it with the step-change only model
pred4b <- predict(model4,type="response",transform(datanew,month=6))/
  mean(data$stdpop)*10^5

{plot(data$rate,type="n",ylim=c(120,300),xlab="Year",ylab="Std rate x 10,000",
  bty="l",xaxt="n")
rect(36,120,60,300,col=grey(0.9),border=F)
points(data$rate,cex=0.7)
axis(1,at=0:5*12,labels=F)
axis(1,at=0:4*12+6,tick=F,labels=2002:2006)
lines(1:600/10,pred3b,col=2) # prediction from the seasonality model
lines(1:600/10,pred4b,col=4) # prediction from the seasonality and change-in-slope model
title("Sicily, 2002-2006")
legend("topleft",c("Step-change only","Step-change + change-in-slope"),lty=1,
  col=c(2,4),inset=0.05,bty="n",cex=0.7)}
```




```
# Test if the change-in-slope improve the fit
# The selected test here is an F-test, which accounts for the overdispersion,
# while in other cases a likelihood ratio or wald test can be applied.
anova(model3,model4,test="F")
```

```
## Analysis of Deviance Table
##
## Model 1: aces ~ offset(log(stdpop)) + smokban + time + harmonic(month,
##     2, 12)
## Model 2: aces ~ offset(log(stdpop)) + smokban * time + harmonic(month,
##     2, 12)
##   Resid. Df Resid. Dev Df Deviance      F Pr(>F)
## 1        52     117.27
## 2        51     115.81  1    1.453 0.6388 0.4279
```

```
# Not surprisingly, the p-value is similar to that of the interaction term
```

Appendix 1: A checklist for conducting DD analysis

This is additional material but you may find it helpful to look at this table when conducting your DD analyses or when reading DD studies.

Ryan et al (2015) provide a helpful checklist of things to look for in a diff-in-diff analysis (copied word-for-word below):

Confirm that	How to test	What to do if violated
1. **Data** exist on study outcomes for at least one observation period among groups exposed and not exposed to an intervention, both **before and after** the intervention was implemented	Directly observable	NA
2. Trends in outcome performance prior to an intervention are **parallel** between treatment and comparison groups	Test equivalence of linear trends between treatment and comparison groups prior to intervention by testing the significance of the interaction term between the time trend and the treatment group	If multiple comparison groups are available, match treatment and comparison units
3. **Baseline outcome levels** are unrelated to expectations for changes in outcomes	For both treatment and comparison groups, test whether baseline outcome is correlated with change in performance across the study period	If multiple comparison groups are available, match treatment and comparison units [The authors recommend propensity score matching.]
4. **Violations to standard statistical assumptions** are appropriately addressed	Test for violations of homoscedasticity of standard errors.	Permutation tests or clustered standard errors will likely result in the most accurate statistical inference when using difference-in-difference analysis
5. Events or factors other than treatment, occurring at the time of treatment, **do not differentially affect outcomes** for treatment and comparison groups	Not directly testable	NA
6. The **composition of treatment and comparison groups does not change** over the course of the study	Test for difference in observed covariates between treatment and comparison rates before and after the intervention. Test for differential drop-out rates between treatment and comparison groups.(Hausman and Wise 1979)	Control for differences in observed covariates between treatment and comparison rates before and after the intervention

Confirm that	How to test	What to do if violated
7. **Treatment does not spill-over** from treatment group to comparison group	Test whether comparison group experiences deviation from existing trend concurrent with intervention	If multiple comparison groups are available, choose alternative comparison group that is not subject to spillovers

Appendix 2: Interpretation of DD model with multiple time points, multiple groups, and different treatment timings

In the main text of this document, we discussed how the following model with group and time fixed effects allows us to estimate the DD estimand in a scenario where we have multiple groups, multiple time points, and different groups receiving the treatment at different time points:

$$Y_{ist} = \beta_s + \gamma_t + \delta D_i + \epsilon_{ist}$$

As a refresher, in the model above, β_s refers to group fixed effects, γ_t refers to time fixed effects, and δ is the DD paramter of interest.

When we have multiple groups, multiple time points, and different treatment start dates, the interpretation of the DD estimand is tricky. A recent paper by Goodman-Bacon (2019) shows that in such a scenario, δ is a weighted average of all possible 2x2 DD comparisons in the study where the weights are proportional to the size of each group and the variance of the treatment dummy for each group (thus, groups that are smaller in size or receive treatment very early or late during the study period receive lower weight).

What are the various 2x2 comparisons? As an example, let's look at the figure below. This figure, which is based on Figure 1 of Goodman-Bacon (2019), shows three groups: the group outlined in red receives treatment "early", the group outlined in blue receives treatment "late", and the final group outlined in green ("control") does not receive treatment over the entire study period. Early treatment starts at $t = 1$ while late treatment starts at $t = 2$.

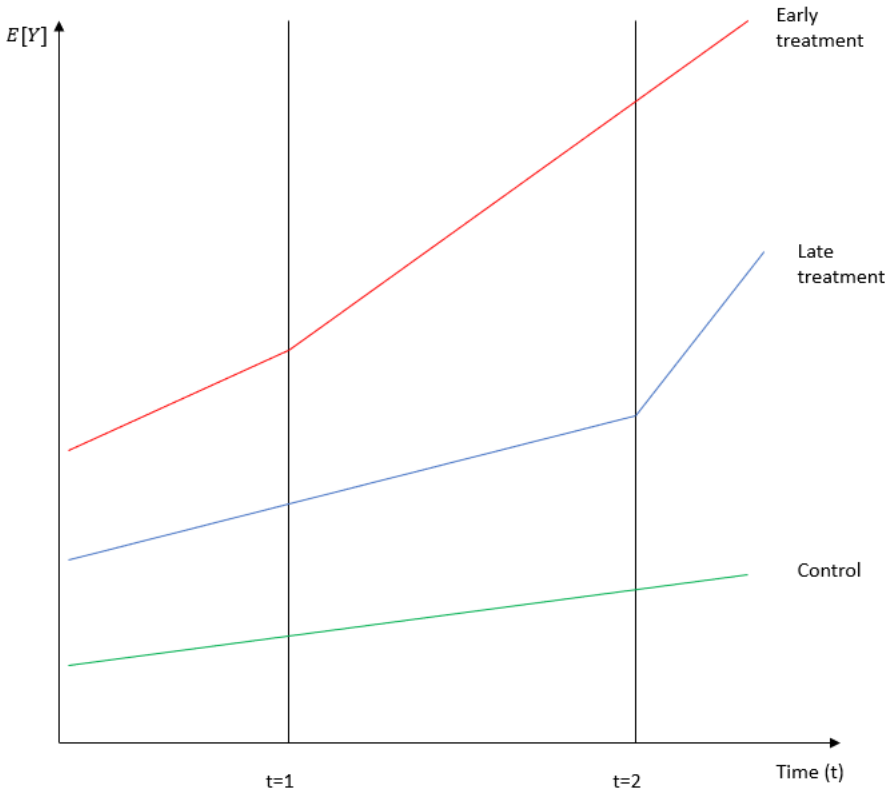


Figure 5: DD analysis with multiple groups and time points

This design leads to four possible 2x2 comparisons between treated and controls:

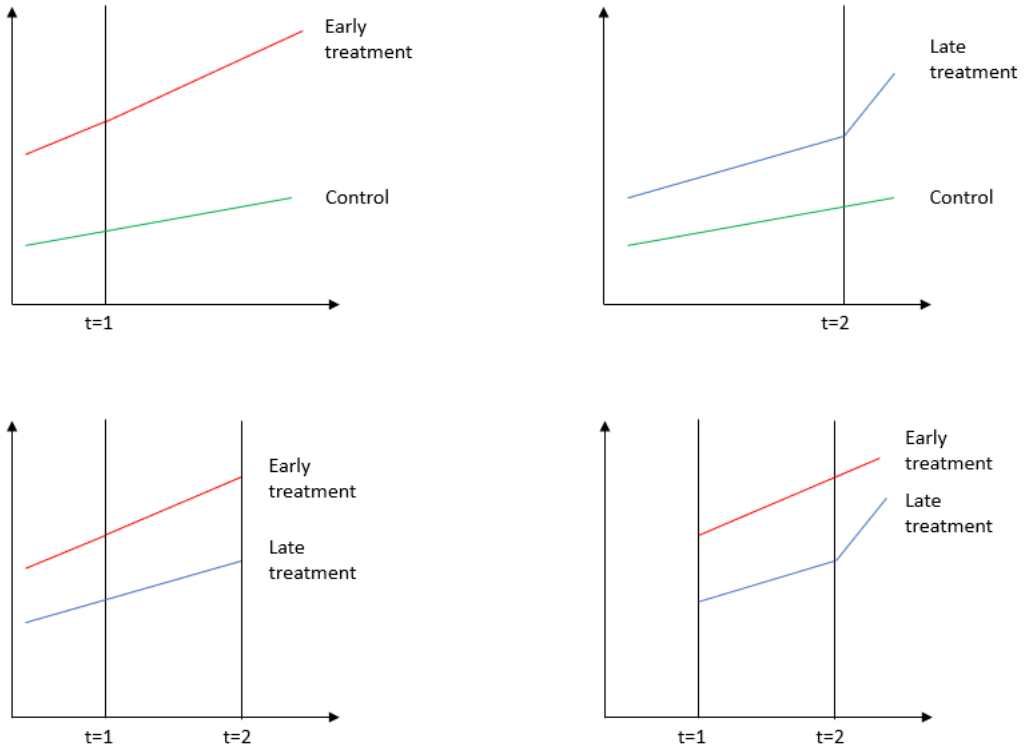


Figure 6: All 2x2 comparisons in a multiple group multiple time point DD analysis

δ is a weighted average of the difference-in-differences estimates across all four of these comparisons. The practical significance of this result is that we are now able to know which comparisons affect our DD estimate most and which comparisons affect it least. Specifically, we know from Goodman-Bacon's result that the DD estimate you get from the regression model above is influenced primarily by those groups that received treatment somewhere close to the middle of the study period and by those groups that are larger in size. This has implications for how generalizable we consider our ATT estimate to be even among the treated (beyond the broader external validity concerns).

Appendix 3: Confounding in a difference-in-difference study

In the main text of this lab, we presented the fact that a random variable C should be considered to be a confounder only if:

1. C is associated with the treatment
2. The relationship between C and the outcome is time-varying **or** the distribution of the covariate across the treated and control groups changes differentially over time

Let's try to see why this is the case. Note that this exposition is drawn directly from Zeldow and Hatfield (2019)

Let Y_{it}^{a} represent the potential outcome for individual i in time period t when the exposure A is set to a. For simplicity, let's assume we are in the 2x2, canonical DD world. This means that we have two groups, one treated and one untreated, and two time periods. Furthermore, let's assume that the treatment effect is additive, as is usually assumed in the Economics literature. Finally, let's assume that we have a binary treatment.

We can write the potential outcomes under $A = 0$

$$Y_{it}^{a=0} = \alpha + \beta_t I(t = 1) + \gamma I(a_i = 1) + \epsilon_{it}$$

Here, $I(t = 1)$ is an indicator variable which equals 1 when we are in the post-treatment time period and 0 otherwise. Similarly, $I(a_i = 1)$ is an indicator variable which equals 1 when we are considering the treatment group and 0 otherwise.

Note that under no treatment, in the treated group,

$$E[Y_{i1}^{a=0} - Y_{i0}^{a=0}|A = 1] = \beta_t$$

and in the control group,

$$E[Y_{i1}^{a=0} - Y_{i0}^{a=0}|A = 0] = \beta_t$$

This suggests that assuming we have correctly specified the model, the parallel trends assumption holds. Second, the calculations we have done above also show us why differences across groups at baseline, as long as they are due to time invariant factors (in our cases, γ) don't matter in a DD analysis since they just get differenced out.

Now, let us introduce a covariate C into the model. Let the potential outcomes under no treatment be:

$$Y_{it}^{a=0} = \alpha + \beta_t I(t = 1) + \gamma I(a_i = 1) + \lambda_t C_i + \epsilon_{it}$$

Here, λ_t represents a potentially time varying effect of C on potential outcome under no treatment. Again, let's consider the difference in potential outcomes under no treatment in the treated and control groups separately between the pre-treatment and post-treatment periods.

In the treated group,

$$E[Y_{i1}^{a=0} - Y_{i1}^{a=0}|A = 1] = \beta_t + \lambda_1 E[C_i|A = 1] - \lambda_0 E[C_i|A = 1]$$

Similarly, in the control group,

$$E[Y_{i1}^{a=0} - Y_{i1}^{a=0}|A = 0] = \beta_t + \lambda_1 E[C_i|A = 0] - \lambda_0 E[C_i|A = 0]$$

Under the parallel trends assumption, we need,

$$\beta_t+\lambda_1 E[C_i|A=1]-\lambda_0 E[C_i|A=1] = \beta_t+\lambda_1 E[C_i|A=0]-\lambda_0 E[C_i|A=0] \Rightarrow \lambda_1 E[C_i|A=1]-\lambda_0 E[C_i|A=1] = \lambda_1 E[C_i|A=0]-$$

Thus, the parallel trends assumption will hold if $\lambda_0 = \lambda_1$ or if $E[C_i|A=0] = E[C_i|A=1]$. As a consequence, when either of these two conditions don't hold and when C is associated with A, then we have confounding.

Appendix 4: Probing the difference-in-differences identifying assumption

A number of studies try to validate the DD identifying assumption by testing for parallel pre-trends. The story goes: if we find statistical evidence to back up the claim that the trends in the outcome in the pre-treatment period between the treated and control groups are parallel, then they will remain parallel without the intervention (and assuming common shocks). There are two issues with this argument:

1) It is not necessarily the case that parallel trends in the pre-treatment period are necessary or sufficient to guarantee parallel trends in the absence of treatment.

2) The standard statistical tests for parallel trends don't allow us to "accept" the claim that the trends in the outcome are indeed parallel. The null hypothesis in these tests is that the trends are parallel; however, as you may recall, in the hypothesis testing framework, we can only ever fail to reject the null and we can never accept it.

Hatfield and colleagues have written papers to suggest that perhaps what may be more useful in testing parallel trends are "equivalence tests". Equivalence tests, as the name suggests, are statistical tests that allow the trends in the outcome in the pre-treatment period between the treated and control groups to deviate. Importantly, however, the amount of deviation is constrained. For a thorough discussion of this, check out this link.

Appendix 5: Multiple time period difference-in-differences

Causal estimand

These days, we usually work with data over multiple time periods. Suppose $T = \{0, ..., T_0, ...T\}$ such that for all $t \in T$, $t \leq T_0$ refers to the pre-treatment period and $t > T_0$ refers to the post-treatment time period. In this case, the causal estimand of interest would be one of the following two depending on our research question:

1. For any $t > T_0$,

$$\delta_{DD} \equiv E[Y_{it}^{a=1} - Y_{it}^{a=0} | A = 1]$$

2. Alternatively, for all $t > T_0$,

$$\delta_{DD} \equiv E[\bar{Y}_{it>T_0}^{a=1} - \bar{Y}_{it>T_0}^{a=0} | A = 1]$$

Note, $\bar{Y}_{it>T_0}^{a}$ refers to the average value of the outcome across all post-treatment time periods under treatment $A = a$.

Identifying assumptions

In the multiple time-period case, one can theoretically adopt different identifying assumptions. For now, we will present the *strictest* identifying assumption in the multiple time-period case. Let $T = \{0, ..., T_0, ...T\}$ such that for all $t \in T$, $t \leq T_0$ represent the pre-treatment period and $t > T_0$ represent the post-treatment period. The multiple time period DD is identified when:

$$E[Y_{it*}^{a=0} - Y_{it'}^{a=0} | A = 1] = E[Y_{it*}^{a=0} - Y_{it'}^{a=0} | A = 0]$$

Where $t* > T_0$ and $t' \leq T_0$. In other words, this identifying assumption forces trends under no treatment to be parallel for every single time point between the pre- and post-treatment periods across the treated and control groups.

Note, once again, that we cannot use data to test this assumption because $Y_{it*}^{a=0}$ is not observable among $A = 1$. Less stricter versions of this assumption posit that the average trends under no treatment are parallel across the treated and control groups between the pre- and post-treatment periods. For a thorough discussion on this, see Zeldow and Hatfield (2019).

Appendix 6: Best practices for visualization

Part of the strength of difference-in-difference estimation and interrupted time series analysis is that the results can be displayed visually. Well-constructed graphs can be a valuable tool for communicating a lot of information very quickly. They help the reader evaluate the visual evidence to support your conclusions and draw inferences from your findings.

In January 2021, a review of best practices for visual representations of interrupted time series analysis was published by a team in Australia and Canada Turner et al. (2021). The recommendations include:

- plotting each individual data point
- using the same data points for visualization as we used in the analysis
- lining up data points with x-axis tick marks
- indicating interruptions with vertical lines or shading with the interruption labeled
- plotting the fitted pre- and post-interruption trends
- indicating the counterfactual trend line

The authors very helpfully illustrate their recommendations by taking previously-published ITS graphs and re-visualizing them to incorporate these recommendations. One such graph, an ITS visualization from a 2017 study of Florida's "Stand Your Ground" law and its effects on homicide and homicide by firearm by David Humphreys and colleagues, is presented and re-visualized by the authors:

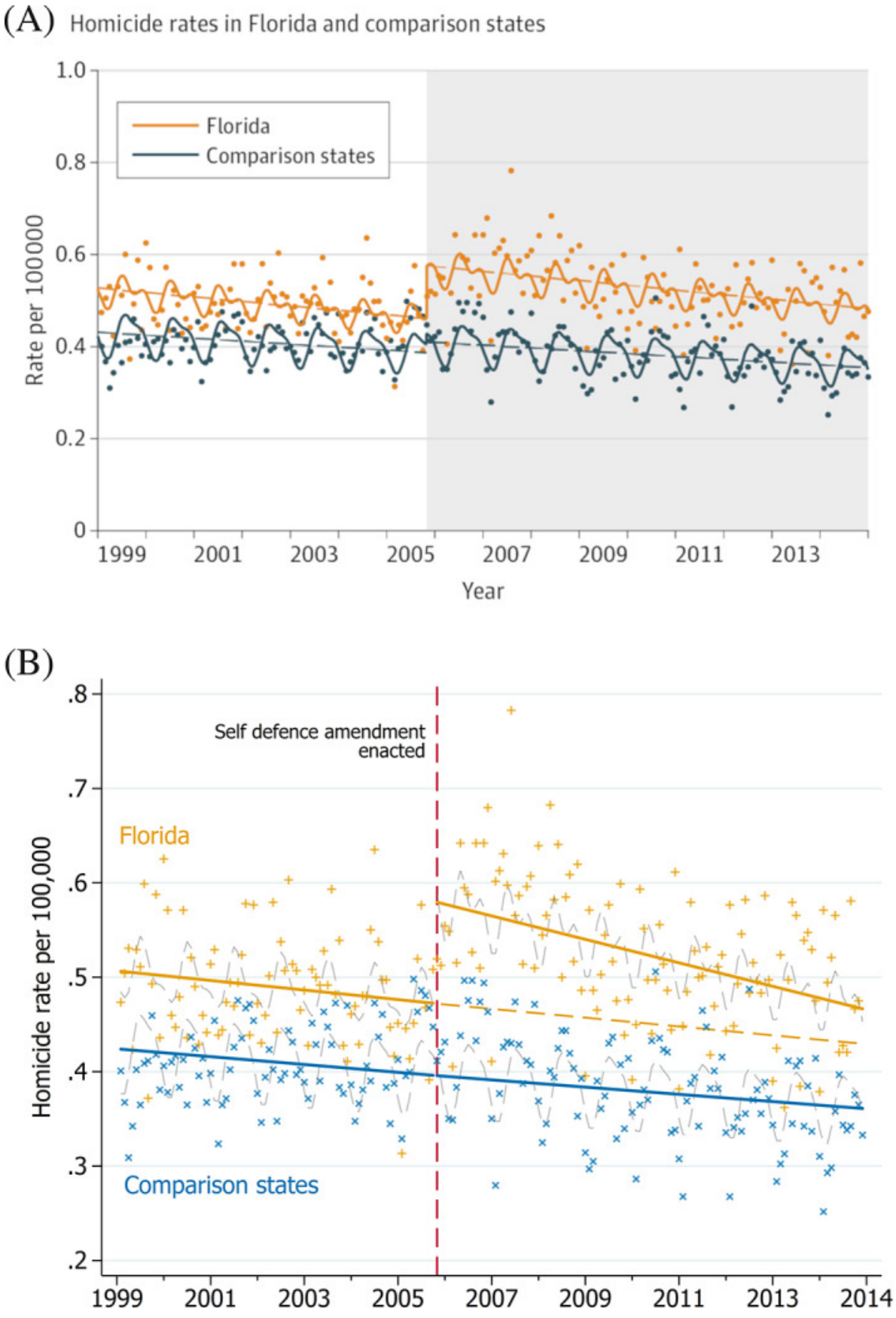


Figure 7: (A) The original visualization published by Humphreys et al. 2017 and (B) the revised version published by Turner et al. 2021