Interrupted Time Series with Aggregated Data: A Review of Methods
Jaffa Romain

Abstract

Interrupted time series is commonly used in place of randomized control trials to evaluate the causal effects of health interventions. Segmented regression is commonly used to analyze interrupted time series data. Many studies perform segmented regression on aggregated data, which has been shown to lead to lack of precision and loss of power. There is a methodological gap in the study of interrupted time series and its analysis. Few studies explore the implications of applying statistical methods such as segmented regression on aggregated data. We compare segmented regression to two alternative methods for analyzing interrupted time series data: weighted segmented regression and generalized mixed effect models segmented regression. The methods are applied to data investigating birth rate with a major sociocultural event as the intervention.

Introduction

Randomized control trials (RCTs) are used to evaluate the causal effects of health interventions in a population. In RCTs, subjects are randomly allocated to either an intervention group or control group and predefined outcomes measured in prospective follow-up are compared between the control and intervention groups. Although RCTs allow for a fair comparison between intervention and control groups, there are cases where implementing a randomized control study is not possible or optimal. By design, RCTs cannot account for variations in health and overall health care complexity (Cruz, 2017). Moreover, RCTs can be costly, unethical and not possible for health policies targeted at the population level (Bonell et al., 2011).

Within the last decade, the use of ITS in health research has almost tripled (Ewusie et al., (2019)) and is generally regarded favorably for evaluating the impact of health interventions and policies. ITS allows for the inference of causality while still accounting for contextual and temporal factors (Cruz, 2017), and unlike RCTs, ITS is especially suitable for population-level interventions over defined and equally spaced time points (Bernal, 2017). ITS designs use repeated observations of an outcome over time. Rather than comparing a control group and intervention group, there is a comparison within a single population pre- and post-intervention. From the pre intervention data, time series of the outcome of interest is taken to estimate an underlying trend. Assuming this trend continues into the post intervention period and continues unchanged, we compare what the trend would look like in the absence of the intervention and the actual trend observed to estimate the effect of the intervention (Pape, 2013).

There is substantial literature on proposed methods for ITS data. Factors including type of outcome (e.g., continuous vs. discrete), data distribution, autocorrelation or seasonality, and the inclusion of a control group are considered when deciding on the method of choice. However, methodological gaps have emerged in current ITS methods. Many studies fail to account for heterogeneity across patients or sites. Applying analytical methods while not accounting for this heterogeneity is associated with increased bias and decreased precision when used on aggregated data (Sen, 1997). Methods used to analyze ITS such as SLR, ARIMA, and segmented regression models do not account for imprecision due to variability of patient outcomes or variability across sites at a certain time point, leading to spurious results. Segmented regression (SR), though commonly used in the analysis of interrupted time series (ITS) data, has been shown to produce spurious results when applied to aggregated data.

In this paper, we present and compare analytical methods used to analyze causal effects of health interventions. We illustrate a comparison of results from analyses using weighted segmented regression and segmented generalized mixed effect models that account for aggregated data to the commonly used segmented regression model. The structure of the article is as follows. In the `Methods` section we look at segmented regression as a method for analyzing ITS data. We then discuss weighted segmented regression and segmented generalized mixed effect models as proposed methods for accounting for aggregated data. These methods are then illustrated and compared by replicating a study conducted by Rodgers, John, and Coleman (2005) that investigated fertility rates in Oklahoma after the 1995 Oklahoma bombing. In the 'results' Section, the findings in this analysis and results from previous studies are discussed. Section 'Discussion' contains some practical recommendations and a brief discussion of the findings and future steps.

Methods

Segmented Regression

Segmented regression is a special case of multiple linear regression with an indicator variable representing the intervention periods, a continuous variable representing the time at which observations are taken, and an interaction variable for the time and intervention period (Kong). A regression line is fitted to each segment of the time series to represent the level and trend pre- and post-intervention. While powerful, segmented regression restricts the interruption to a predetermined time point in the series, neglecting the plausible differences in autocorrelation and variability in the data (Cruz, 2017).

For segmented regression the following model is used:

$Y_t = \beta_0 + \beta_1 t + \beta_2 X + \beta_3 Xt$
where:
t = time elapsed since the start of the study
X_t = indicator variable for time period (0 = pre-intervention, 1 = post-intervention)
Y_t = outcome of interest at time t

β_0 = baseline level at T = 0
β_1 = change in outcome associated with unit increase in time
β_2 = change in outcome level following intervention
β_3 = interaction between time and intervention

Weighted Segmented Regression

Ewusie et al. (2019), propose an extended segmented regression approach, weighted segmented regression, to account for between participant variation and differences across multiple sites. The segmented regression approach assumes homoscedasticity. That is, $var(\epsilon_i) = \sigma^2$. Specifically, the variance associated with the data at each time point is constant and due to random error. However, ITS studies have aggregated data at a given point in time. This assigns some level of statistical uncertainty. Thus, the variability of the data is actually not constant, but rather, v$var(y_i) = var(\epsilon_i) = \sigma_i^2, and\ \sigma_i^2 = \sigma^2 * \gamma_i^2$ where γ_i^2 is a function of heterogeneity in patients and the number of patients. Consequently, applying the method of least squares will not be optimal for estimation since we do not get minim variance unbiased estimators. Weighted segmented regression addresses the limitation of the OLS approach using weighted least squares. The data are weighted to account for differences in variance. We derive the estimators in a weighted segmented regression framework through minimizing the weighted sum of squares. In this matrix from this gives:

$$\sum w_i \epsilon_i^2 = (Y - X\beta)' W (Y - X\beta) = Y'WY - Y'WX\beta - \beta X'WY + \beta' X' WX\beta$$

The model estimators are then given by: $\widehat{\beta_w} = (X'WX)^{-1}X'WY$
Where W represents a diagonal matrix with the vector w = $(w_1, w_2, \ldots, w_k)$.

Segmented generalized mixed effect Model

According to French and Heagerty (2008), generalized mixed models may be used in combination with segmented regression for an ITS design (French). Where the β's are the fixed effects, v_{i0} is the random intercept and v_{i1} is the random slope. The random effects account for repeated measures for the ith individual or hierarchical groupings such as site or hospital. The random effects reflect the heterogeneity across these repeated measures or groupings allowing for efficient group-specific intercepts. The generalized mixed model allows for less rigid assumptions to produce consistent estimates. However, they are not statistically efficient, as they require the estimation of separate parameters at each cluster (Strumpf et al. 2017) .
The fitted segmented generalized mixed effect model is of the form:

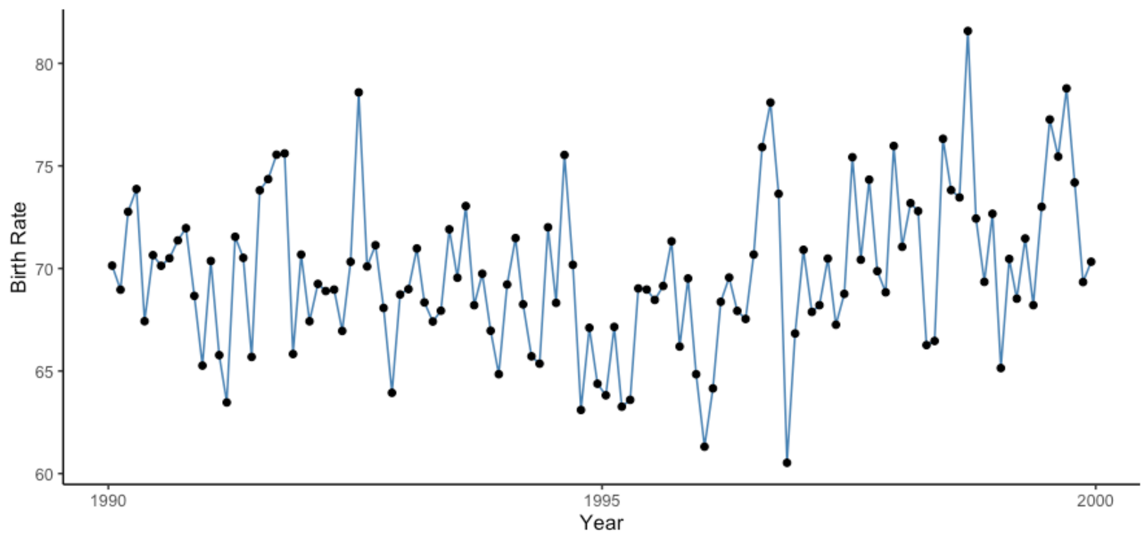
$$Y_t = (\beta_0 + v_{i0}) + (\beta_1 + v_{i1})t + \beta_2 X + \beta_3 Xt$$

Results

Illustrative example: Fertility rates in Oklahoma after the Oklahoma City Bombing

The Oklahoma City bombing was a domestic terrorist attack occurring on April 19, 1995. There were 168 fatalities and 800 injuries. Replacement theory, community influence theory, and terror management. suggest that political and sociocultural events can affect fertility (Rodgers et al., 2005). Here we illustrate the three methods discussed above using the openly available data 12 different counties. This is intended to only be a simple illustration of methods; results should be interpreted with caution. The data consists of 1440 observations. The data are aggregated across the 12 counties, providing birth rates summarized across counties at a given time point (see figure 1).

Figure 1: Birth rates for 12 Oklahoma Counties from 1990 -2000.



For this study, 1440 observations across 12 counties, with 756 observations pre-intervention and 684 post intervention. For analysis of the birth rate of across all 12 counties, estimates of the regression coefficients were obtained for the segmented regression, weighted segmented regression, and the generalized linear mixed effect model with segmented regression. All the models produced a negative coefficient for the time variable, indicating a decline in birthrate prior to the bombing. For segmented regression and weighted segmented regression, the positive coefficient suggests that there was in immediate increase in the percentage of birth rates during the occurrence of the intervention replicating the findings of the original study conducted by Rodgers et al., (2005). Segmented regression and weighted segmented regression produced relatively comparable estimates, while the mixed model method differed greatly from the other methods. Further, the mixed effects method had the widest confidence interval, while the segmented regression estimate appeared to have the narrowest confidence interval.

Table 1: Estimates for Intervention Effects and 95% Confidence Intervals for 3 Methods

Variable	Segmented Regression Estimate (95% CI)	Weighted Segmented Regression (95% CI)	Generalized linear mixed effects segmented regression
β_0 (Baseline level)	58.37 (56.94, 59.80)	52.762 (48.42, 57.104)	63.90 (52.53, 73.95)
β_1 (time)	-1.023 *** (-2.30, 0.26)	-1.46 ** (-2.76, -0.16)	-0.0029 (-0.015, 0.011)
β_2(Difference in level following intervention)	0.0043 (0.003, 0.006)	0.077 * (0.02, 0.134)	-4.45 ** (-7.7, -1.8)
β_3 = interaction between time and intervention	-0.077 (-0.13, -0.024)	-0.078 (-0.14, -0.021)	-0.0011 (-0.0043, -0.0038)

Note. P-value < 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Discussion

In this study, we aimed to extend on the existing literature of interrupted time series by exploring alternative methods to the traditional segmented regression approach to address the issues that come with working with aggregated data. Although segmented regression is the most used framework, variability due to the aggregated nature of the data is overlooked when modelling. we investigated methods that considered heterogeneity across groups. We performed an analysis using the data looking into birthrates in Oklahoma counties to compare segmented regression to weighted segmented regression and generalized mixed effects with segmented regression. Ewusie et al., (2020) conducted a previous analysis investigating overall patient mobility across 14 hospitals. Their results showed that among the three methods segmented regression, pooled analysis, and weighted segmented regression, weighted segmented regression had the narrowest confidence interval. Conversely, the traditional segmented regression had the widest confidence interval. Similar to their study, the parameter estimates of our models were comparable for the two segmented regression methods.

We found that segmented regression and weighted segmented regression produced similar values, however contrary to findings of Ewusie et al., (2020), we did not find that weight segmented regression performed much better than the standard segmented regression. In some cases, such as the estimate for level differences following intervention, segmented regression appears to a narrower confidence interval. Overall, the estimates for all models gave consistent conclusions. For example, time was estimates to have a negative coefficient estimate for all 3 models. However, it was found to only be statistically significant in the segmented and weighted segmented regression approaches.

Further analysis is needed before making any concrete conclusions. Specifically, more extensive simulations are needed to test the different methods under different scenarios. First, the methods should be compared under different scenarios that consider ITS data from different data distributions. For example, skewed ITS data could lead us to very different conclusions, compared to normally distributed data. Another limitation was the data set was only aggregated at one level. We did not consider a scenario where there is aggregation at two levels: site (location, hospital, etc.) and patient level. Few studies compare the generalized mixed effect model to segmented models. A major strength of the mixed effect approach is that by having repeated measures on the same participant before and after exposure, we can control time-invariant confounders (Saeed, 2018). However, time-varying confounders can be present and may bias results (Saeed, 2018).

Conclusion

In summary, comparing segmented regression to weighted segmented regression and generalized mixed effect models showed that the models provided similar estimates, but notable differences in variance. The generalized mixed effect model appears to have the highest standard error among the 3 models. Although our study does not show that these the models accounting for aggregation performed better, we have low generalizability by comparing the methods on a single data set. However, the difference in variance among the methods emphasize the need to further explore ways to account for differences in participant populations and settings

References

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. https://doi.org/10.1093/ije/dyw098

Bonell CP, Hargreaves J, Cousens S, Ross D, Hayes R, Petticrew M, Kirkwood BR. Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions. J Epidemiol Community Health. 2011 Jul;65(7):582-7. doi: 10.1136/jech.2008.082602. Epub 2009 Feb 12. PMID: 19213758.

Cook TD, Campbell DT. A. Quasi-experimentation: Design & analysis issues for field settings. Boston: Houghton Mifflin; 1979.

Cruz M, Bender M, Ombao H. A robust interrupted time series model for analyzing complex health care intervention data. Stat Med. 2017 Dec 20;36(29):4660-4676. doi: 10.1002/sim.7443. Epub 2017 Aug 29. PMID: 28850683.

Ewusie J. E., Soobiah C, Blondal E, Beyene J, Thabane L, Hamid JS. Methods, Applications and Challenges in the Analysis of Interrupted Time Series Data: A Scoping Review. J *Multidiscip Healthc*. 2020;13:411-423
https://doi.org/10.2147/JMDH.S241085

Ewusie, J. E., Thabane, L., Beyene, J., Straus, S. E., & Hamid, J. S. (2019). MultiCenter Interrupted Time Series Analysis: Incorporating Within and Between-Center Heterogeneity. *Clinical Epidemiology*, *12*, 625-636. https://doi.org/10.2147/CLEP.S231843

French B, Heagerty PJ (2008) Analysis of longitudinal data to evaluate a policy change. Stat Med 27:5005–5025. https://doi.org/10.1002/sim.3340

James Lopez Bernal, Steven Cummins, Antonio Gasparrini, Interrupted time series regression for the evaluation of public health interventions: a tutorial, *International Journal of Epidemiology*, Volume 46, Issue 1, February 2017, Pages 348–355, https://doi.org/10.1093/ije/dyw098

Kong M, Cambon A, Smith MJ. Extended logistic regression model for studies with interrupted events, seasonal trend, and serial correlation. *Commun Stat Theory Meth*. 2012;**41**(19):3528–3543. doi: 10.1080/03610926.2011.563020)

Muggeo, Vito. (2008). Segmented: An R Package to Fit Regression Models With Broken-Line Relationships. R News. 8. 20-25.

Pape, U. J., Millett, C., Lee, J. T., Car, J., & Majeed, A. (2013). Disentangling secular trends and policy impacts in health studies: use of interrupted time series analysis. *Journal of the Royal Society of Medicine*, *106*(4), 124-129. https://doi.org/10.1258/jrsm.2012.110319\

Rodgers JL, St John CA, Coleman R. Did fertility go up after the Oklahoma City bombing? An analysis of births in metropolitan counties in Oklahoma, 1990-1999. Demography. 2005 Nov;42(4):675-92. doi: 10.1353/dem.2005.0034. PMID: 16463916.

Saeed S, Moodie EEM, Strumpf EC, Klein MB. Segmented generalized mixed effect models to evaluate health outcomes. Int J Public Health. 2018 May;63(4):547-551. doi: 10.1007/s00038-018-1091-9. Epub 2018 Mar 16. PMID: 29549396.

Sen, A., & Srivastava, M. (1997). Regression analysis: theory, methods, and applications. Springer Science & Business Media.

Strumpf EC, Harper S, Kaufman JS (2017) Fixed effects and difference-in-differences. In: Methods in social epidemiology, 2nd edn. Jossey-Bass, San Francisco, pp 341–368

Wang JJ, et al. A comparison of statistical methods in interrupted time series analysis to estimate an intervention effect. Australasian Road Safety Research, Policing and Education Conference 2013

Yelland, J., Riggs, E., Szwarc, J. et al. Bridging the Gap: using an interrupted time series design to evaluate systems reform addressing refugee maternal and child health inequalities. *Implementation Sci* 10, 62 (2015). https://doi.org/10.1186/s13012-015-0251-z

Appendix

```
dat$time <- rep( 1 : nrow( dat ))
dat$treatment <- ifelse( dat$Year > 1995 | dat$Year == 1995 & dat$Month >= 4, 1, 0)
table(dat$treatment)
dat$timeSince <- c(rep(0, 64), rep(1:56))
# dat %>% ggplot(aes(x = Date, y = BirthRate)) +
#   geom_line( color="steelblue") +
#   geom_point() + geom_vline(xintercept = 1995 ,color = "red" ) +
#   xlab("") +
#   theme(axis.text.x=element_text(angle=60, hjust=1)) + theme_classic() +
#   labs(y = "Birth Rate", x = "Year")

library(segmented)
```

```
dat %>% summary()
dat %>% ggplot(aes(x = BirthRate)) + geom_histogram(bins = 30, color = "black", fill = "white") +
theme_classic() + labs(x = "Birth Rate")
# standard segmented regression
mod1 <- lm(BirthRate~ treatment + time, data=dat)
segmodel <- segmented(lm(BirthRate~ treatment + time, data = dat),seg.Z = ~time)
summary(segmodel)
# When not providing estimates for the breakpoints "psi = NA" can be used.
# The number of breakpoints that will show up is not defined
#my.seg <- segmented(my.lm,
#                seg.Z = ~ DistanceMeters,
#                psi = NA)

# display the summary
slope(segmodel)

# weighted segmented regression
# define weights
mod2 <- lm(BirthRate~treatment + time + CountyName, data=dat)
wt <- 1 / lm(abs(mod2$residuals) ~ mod2$fitted.values)$fitted.values^2
segmodel_weighted <- segmented(lm(BirthRate~treatment + time,
                            weights = wt, data = dat), seg.Z = ~time,)

summary(segmodel_weighted)
slope(segmodel_weighted)

# random mixed model
modellme <-lme(BirthRate~treatment*time, random = ~1|CountyName, data=dat)

segmodel_randomeffects <- segmented(modellme,
                             seg.Z = ~time, x.diff = ~treatment,
                           random=list(CountyName=pdDiag(~1+time+U+G0)))
# confidence intervals
summary(segmodel_randomeffects)
# segmented model
sdsegmodel <- sqrt(diag(vcov(segmodel)))
coefsegmodel <- segmodel$coefficients
Upper <- round(coefsegmodel + (1.96 * sdsegmodel), 3)
lower <-  round(coefsegmodel -  (1.96 * sdsegmodel), 3)
segmodel_summary <- cbind(coefsegmodel, lower, Upper)


# weighted segmented
```

```
sdsegmodel_weighted <- sqrt(diag(vcov(segmodel_weighted)))
coefsegmodel_weighted <- segmodel_weighted$coefficients
Upper <- round(coefsegmodel_weighted + (1.96 * sdsegmodel_weighted), 3)
lower <-  round(coefsegmodel_weighted -  (1.96 * sdsegmodel_weighted), 3)
segmodel_summary_weighted <- cbind(coefsegmodel_weighted, lower, Upper)
# mixed model

sdsegmodel_randomeffects <- sqrt(diag(vcov(segmodel_randomeffects)))
coefsegmodel_randomeffects <- segmodel_randomeffects$coefficients
Upper <- round(coefsegmodel_randomeffects + (1.96 * sdsegmodel_randomeffects), 3)
lower <-  round(coefsegmodel_randomeffects -  (1.96 * sdsegmodel_randomeffects), 3)
segmodel_summary_randomeffects <- cbind(coefsegmodel_randomeffects, lower, Upper)

ci_random <- confint(segmodel_randomeffects)
```