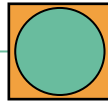


# Video Tutorial: Interrupted Time Series

## Part 2

Data for Impact

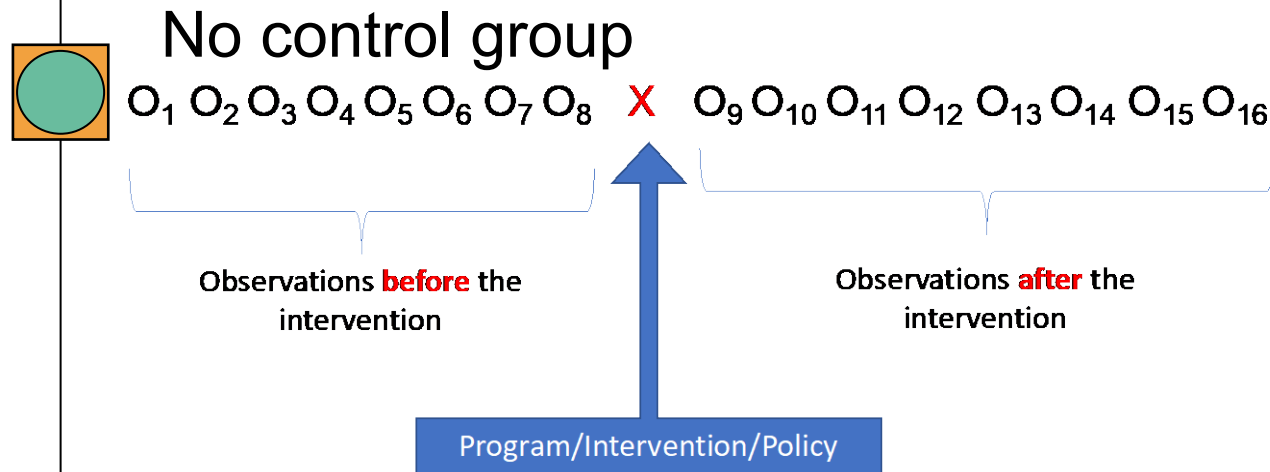




# Overview

- Part 1:
  - Interrupted time series overview
  - Important considerations
  - Preparing the dataset
- Part 2:
  - Interrupted time series analysis: Single group
- Part 3:
  - Interrupted time series analysis: Multiple group

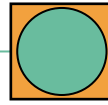
# SINGLE GROUP ITS



## Legend

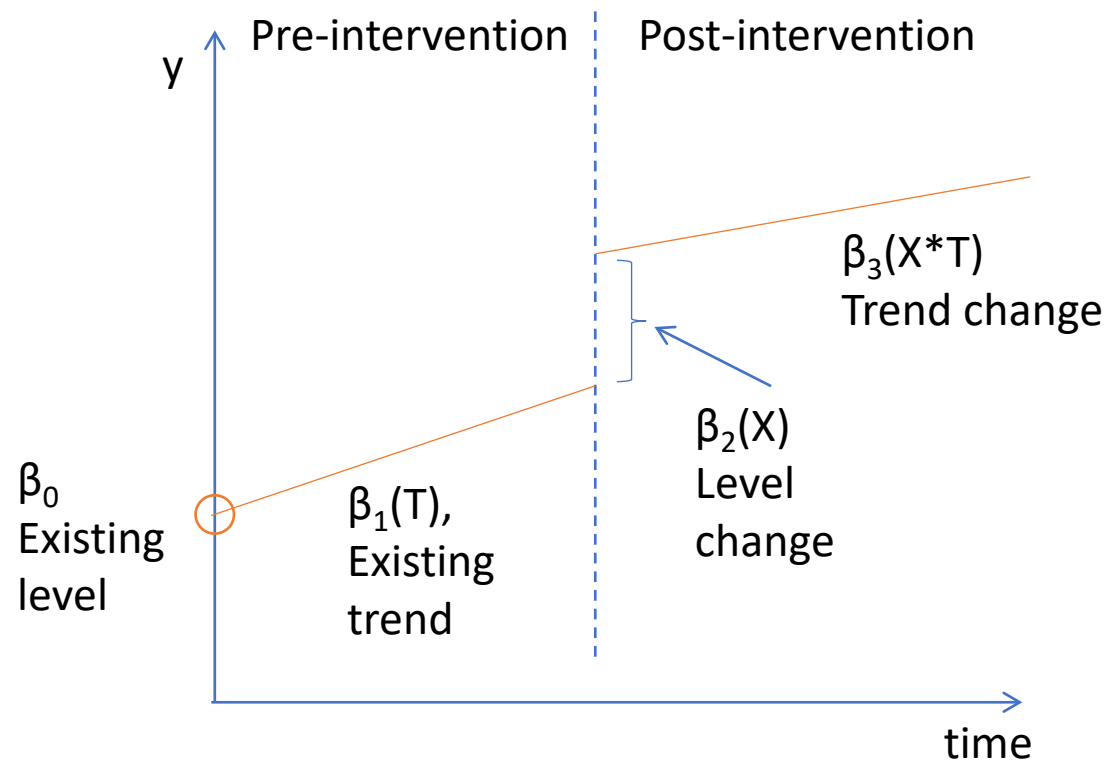
O - Observation

X – Program/Intervention/Policy



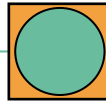
# Single group ITS

$$y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t + \varepsilon_t$$

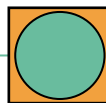


Where,

- $T$  = time
- $X$  = program/intervention/policy
- $XT$  = time after the intervention
- $\varepsilon$  = error term
- $\beta_0$  = Intercept, starting or existing level,
- $\beta_1$  = Trend, trajectory / slope until intervention
- $\beta_2$  = Effect of intervention, level change after intervention
- $\beta_3$  = Trend change
- NB:  $\beta_2$  is the immediate treatment effect and  $\beta_3$  is the treatment effect over time



Year	Rate	T	X	XT
1999	30.001	1	0	0
2000	27.921	2	0	0
2001	20.739	3	0	0
2002	26.409	4	0	0
2003	30.834	5	0	0
2004	27.462	6	0	0
2005	25.635	7	0	0
2006	20.416	8	0	0
2007	29.888	9	0	0
2008	25.051	10	0	0
2009	29.559	11	1	1
2010	30.220	12	1	2
2011	26.707	13	1	3
2012	26.294	14	1	4
2013	23.768	15	1	5
2014	25.714	16	1	6
2015	30.057	17	1	7
2016	24.028	18	1	8
2017	27.309	19	1	9
2018	27.247	20	1	10



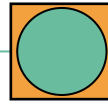
# Sample dataset

- Sample data is from a fictional country called Exile.
- In November 2017, Exile introduced a policy that eliminated fees at health facilities associated with maternity care services, with the aim of increasing utilization of health services.
- It was piloted in two constituencies, Gold-Coast IX and Gold-Coast XV. The policy was not piloted in North Troy XXVI.
- Dataset: Health facility data on the number of women who attended their 4<sup>th</sup> antenatal care (ANC) visits from 2017 to 2020

## Dataset description

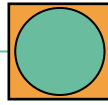
obs: 132  
vars: 6  
size: 3,168

variable name	storage type	display format	value label	variable label
id	float	%9.0g		Unique ID
region	long	%26.0g	de_region	Region
constituency	long	%35.0g	de_constituency	Constituency
nfac_cons	float	%9.0g		Number of health facilities in each constituency
mdate	float	%tm		Date
conANC4	float	%9.0g		4th ANC visit (constituency level)



# Sample dataset

id	region	constituency	nfac_cons	mdate	conANC4
20	Gold Coast Region	Gold-C. IX	17	2017m1	.0012877
20	Gold Coast Region	Gold-C. IX	17	2017m2	.0013797
20	Gold Coast Region	Gold-C. IX	17	2017m3	.0012877
20	Gold Coast Region	Gold-C. IX	17	2017m4	.0013797
20	Gold Coast Region	Gold-C. IX	17	2017m5	.0016557
20	Gold Coast Region	Gold-C. IX	17	2017m6	.0015637
20	Gold Coast Region	Gold-C. IX	17	2017m7	.0016557
20	Gold Coast Region	Gold-C. IX	17	2017m8	.0014717
20	Gold Coast Region	Gold-C. IX	17	2017m9	.0015637
20	Gold Coast Region	Gold-C. IX	17	2017m10	.0015637
20	Gold Coast Region	Gold-C. IX	17	2017m11	.0016557
20	Gold Coast Region	Gold-C. IX	17	2017m12	.0027594
20	Gold Coast Region	Gold-C. IX	17	2018m1	.0039144
20	Gold Coast Region	Gold-C. IX	17	2018m2	.0040924
20	Gold Coast Region	Gold-C. IX	17	2018m3	.0047151
20	Gold Coast Region	Gold-C. IX	17	2018m4	.0056047
20	Gold Coast Region	Gold-C. IX	17	2018m5	.0047151
20	Gold Coast Region	Gold-C. IX	17	2018m6	.0056047
20	Gold Coast Region	Gold-C. IX	17	2018m7	.0055158
20	Gold Coast Region	Gold-C. IX	17	2018m8	.0072951
20	Gold Coast Region	Gold-C. IX	17	2018m9	.0070282
20	Gold Coast Region	Gold-C. IX	17	2018m10	.0065834
20	Gold Coast Region	Gold-C. IX	17	2018m11	.0054268
20	Gold Coast Region	Gold-C. IX	17	2018m12	.0072061
20	Gold Coast Region	Gold-C. IX	17	2019m1	.0111295
20	Gold Coast Region	Gold-C. IX	17	2019m2	.0095153
20	Gold Coast Region	Gold-C. IX	17	2019m3	.0072214
20	Gold Coast Region	Gold-C. IX	17	2019m4	.0103648
20	Gold Coast Region	Gold-C. IX	17	2019m5	.0088356
20	Gold Coast Region	Gold-C. IX	17	2019m6	.01011
20	Gold Coast Region	Gold-C. IX	17	2019m7	.0097701
20	Gold Coast Region	Gold-C. IX	17	2019m8	.0086657

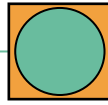


# Research question

## Research Questions

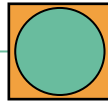
1. What was the impact of the **policy** on the **utilization of ANC, specifically the 4<sup>th</sup> ANC visit of pregnant women**, in Gold Coast IX?
2. What was the impact of the policy on the utilization of ANC, specifically the 4th ANC visit of pregnant women, in the intervention constituency, Gold Coast IX, compared to the control constituency, North Troy XXVI?





## Steps needed to use ITS

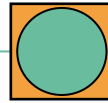
1. Determine time period(s)
2. Select the analytic cohort, if applicable
3. Determine outcomes of interest
4. Setup the data and visually inspect data
5. Preliminary analysis
6. Address autocorrelation, if applicable
7. Run final model
8. Plot results
9. Predict relative and absolute effects



# Determine time period

$$y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t + \varepsilon_t$$

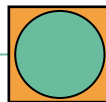
- **Y** : outcome, constituency rate of 4<sup>th</sup> ANC visits
- **T** : time since the start of the study
  - Each month is an observation, count the number of months since the start of the study (Jan 2017)
- **X** : dummy variable indicating pre and post policy period
  - Pre – period (Jan. – Oct. 2017) = 0
  - Post – period (Nov. 2017 – Aug. 2020) = 1
- **XT** : Time after the policy, number of months after start of policy (Nov. 2017). Pre-period=0



# Variables

- T – **time**, variable counts the number of observations
- X – **period**, dummy variable of pre and post policy period
- XT – **posttime**, variable counts the number of observations after the introduction of the policy

mdate	time	period	posttime
2017m1	1	0	0
2017m2	2	0	0
2017m3	3	0	0
2017m4	4	0	0
2017m5	5	0	0
2017m6	6	0	0
2017m7	7	0	0
2017m8	8	0	0
2017m9	9	0	0
2017m10	10	0	0
2017m11	11	1	0
2017m12	12	1	1
2018m1	13	1	2
2018m2	14	1	3
2018m3	15	1	4
2018m4	16	1	5
2018m5	17	1	6
2018m6	18	1	7
2018m7	19	1	8
2018m8	20	1	9
2018m9	21	1	10



# STATA Commands

## Time

```
sort constituency mdate  
  
by constituency : gen time=_n  
  
label var time "Time points"
```

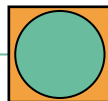
## Period

```
. gen period=0 if time>=1 & time<=10  
(102 missing values generated)  
  
. replace period=1 if time>=11 & time<=45  
(102 real changes made)
```

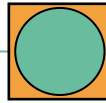
Pre and post policy period	Freq.	Percent	Cum.
Pre-period	30	22.73	22.73
post-period	102	77.27	100.00
Total	132	100.00	

## Posttime

```
. bysort constituency period : gen posttime=_n-1  
  
. replace posttime=0 if period==0  
(27 real changes made)  
  
. label var posttime "Time points after the policy"
```



id	region	constituency	nfac_cons	mdate	conANC4	time	period	posttime
20	Gold Coast Region	Gold-C. IX	17	2017m1	.0012877	1	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m2	.0013797	2	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m3	.0012877	3	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m4	.0013797	4	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m5	.0016557	5	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m6	.0015637	6	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m7	.0016557	7	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m8	.0014717	8	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m9	.0015637	9	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m10	.0015637	10	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m11	.0016557	11	Pre-period	0
20	Gold Coast Region	Gold-C. IX	17	2017m12	.0027594	12	post-period	1
20	Gold Coast Region	Gold-C. IX	17	2018m1	.0039144	13	post-period	2
20	Gold Coast Region	Gold-C. IX	17	2018m2	.0040924	14	post-period	3
20	Gold Coast Region	Gold-C. IX	17	2018m3	.0047151	15	post-period	4
20	Gold Coast Region	Gold-C. IX	17	2018m4	.0056047	16	post-period	5
20	Gold Coast Region	Gold-C. IX	17	2018m5	.0047151	17	post-period	6
20	Gold Coast Region	Gold-C. IX	17	2018m6	.0056047	18	post-period	7
20	Gold Coast Region	Gold-C. IX	17	2018m7	.0055158	19	post-period	8
20	Gold Coast Region	Gold-C. IX	17	2018m8	.0072951	20	post-period	9
20	Gold Coast Region	Gold-C. IX	17	2018m9	.0070282	21	post-period	10
20	Gold Coast Region	Gold-C. IX	17	2018m10	.0065834	22	post-period	11
20	Gold Coast Region	Gold-C. IX	17	2018m11	.0054268	23	post-period	12
20	Gold Coast Region	Gold-C. IX	17	2018m12	.0072061	24	post-period	13
20	Gold Coast Region	Gold-C. IX	17	2019m1	.0111295	25	post-period	14
20	Gold Coast Region	Gold-C. IX	17	2019m2	.0095153	26	post-period	15
20	Gold Coast Region	Gold-C. IX	17	2019m3	.0072214	27	post-period	16



# Declare data

- Declare data to be time series data

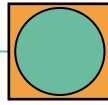
- tsset timevar

```
. tab constituency
```

Constituency	Freq.	Percent	Cum.
Gold-C. IX	44	100.00	100.00
Total	44	100.00	

- tsset panelvar timevar

Constituency	Freq.	Percent	Cum.
Gold-C. IX	44	33.33	33.33
Gold-C. XV	44	33.33	66.67
N-Troy XXVI	44	33.33	100.00
Total	132	100.00	

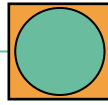


# Declare data

## With panel variable

```
tsset constituency mdate  
  panel variable:  constituency (strongly balanced)  
  time variable:  mdate, 2017m1 to 2020m8  
                delta:  1 month
```

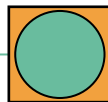
```
tsset constituency time  
  panel variable:  constituency (strongly balanced)  
  time variable:  time, 1 to 44  
                delta:  1 unit
```



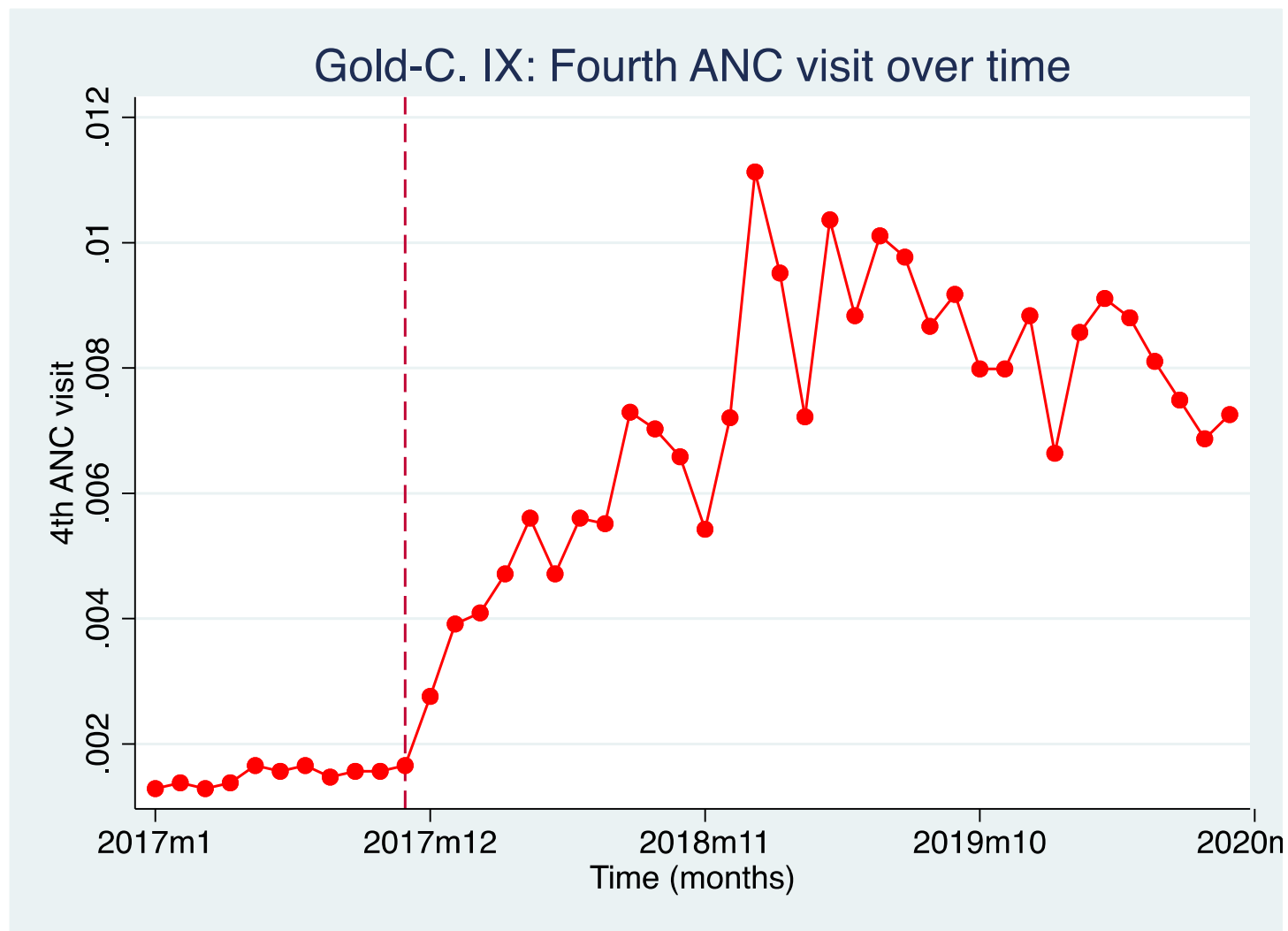
# Visual Inspection

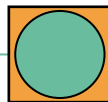
- Graph data of interest
  - STATA command:
    - `twoway connected depvar invar, xline(year)`
    - `twoway line depvar invar, xline(year)`





**STATA command:** `twoway connected conANC4 time if constituency==20, xline(11, lpattern(dash)) lcolor(red) mcolor(red) xtitle(Year (months)) ytitle(1st ANC visit) title(First ANC visit over time)`





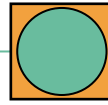
# ITSA command

- Used for single and multiple group ITS

```
itsa depvar [indepvars] [if] [weight], trperiod(numlist) [single treatid(#)  
    lag(#) figure[(twoway_options)] replace
```

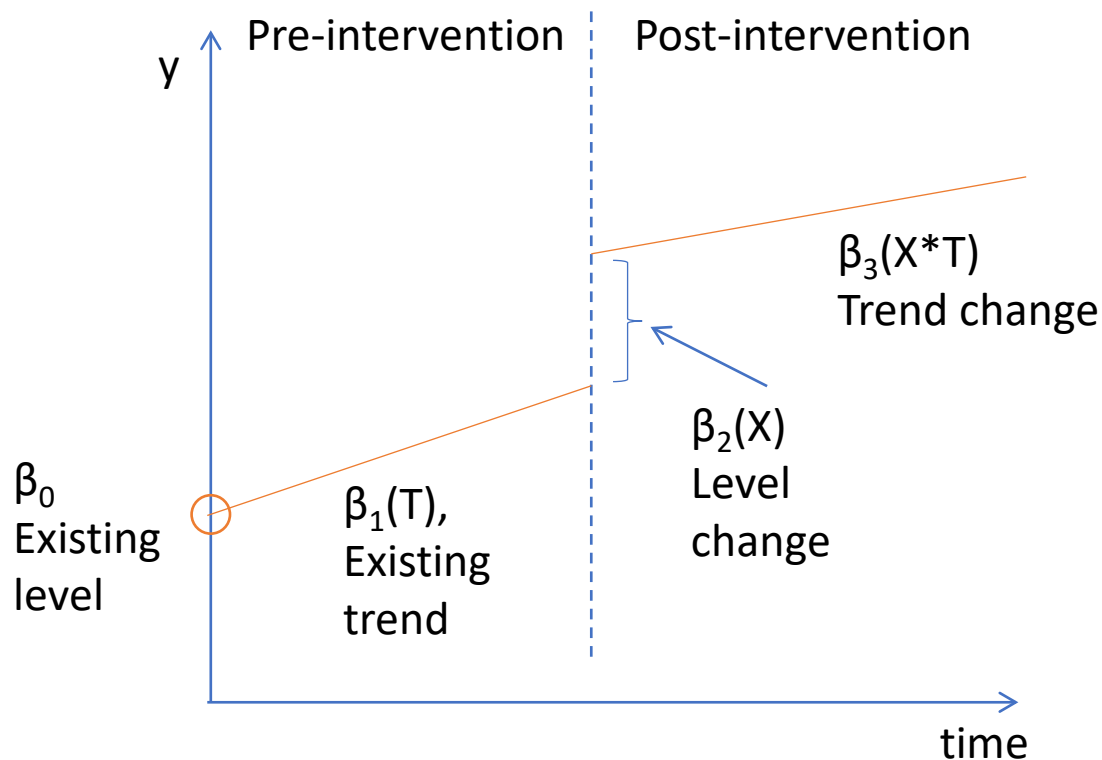
Component	Description
trperiod(numlist)*	Time period when the intervention begins
single	Included for single group analysis.
treatid(#)	Identifier of treatment cohort
lag(#)	# of lags to be considered in the autocorrelation structure
figure[(twoway_options)]	Produces a line plot of predicted dependent variable combined with a scatter plot
replace	Replaces variables created if they already exist

**Tip** : search "help itsa" for more information on the command



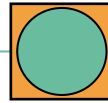
# Reminder: Single group ITS

$$y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t + \varepsilon_t$$



Where,

- $T$  = time
- $X$  = program/intervention/policy
- $XT$  = time after the intervention
- $\varepsilon$  = error term
- $\beta_0$  = Intercept, starting or existing level,
- $\beta_1$  = Trend, trajectory / slope until intervention
- $\beta_2$  = Effect of intervention, level change after intervention
- $\beta_3$  = Trend change
- NB:  $\beta_2$  is the immediate treatment effect and  $\beta_3$  is the treatment effect over time

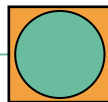


# Preliminary Model

**STATA command:** itsa conANC4, single treatid(20) trperiod(11) lag(0) figure

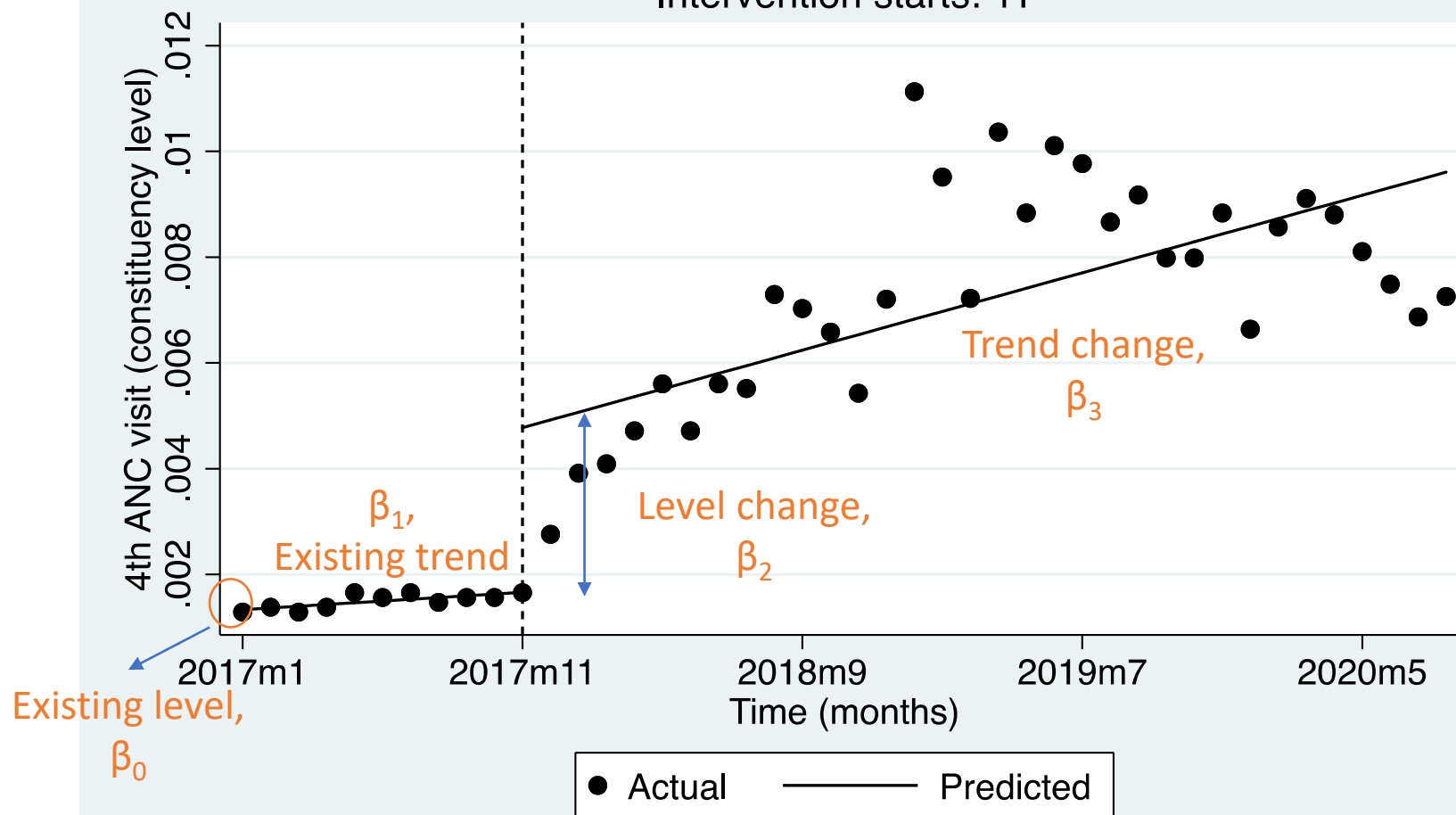
Regression with Newey-West standard errors				Number of obs	=	44
maximum lag: 0				F( 3, 40)	=	137.86
				Prob > F	=	0.0000
conANC4	Newey-West					
	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
_t	.0000329	7.23e-06	4.55	0.000	.0000183	.0000475
_x11	.0031145	.0005963	5.22	0.000	.0019094	.0043196
_x_t11	.0001136	.0000317	3.58	0.001	.0000494	.0001777
_cons	.0013329	.0000441	30.24	0.000	.0012438	.001422

Variable	Explanation	Coefficient	P> t	[95% Conf. Interval]
_t	Existing trend ( $\beta_1$ )	0.0000329	0.000	[0.0000183, 0.0000475]
_x11	Level change ( $\beta_2$ )	0.0031145	0.000	[0.0019094, 0.0043196]
_x_t_11	Trend change ( $\beta_3$ )	0.0001136	0.001	[0.0000494, 0.0001777]
_cons	Existing level ( $\beta_0$ )	0.0013329	0.000	[0.0012438, 0.0014220]

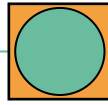


## ANC (4th) visits over time

Intervention starts: 11

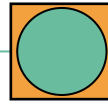


Regression with Newey-West standard errors - lag(0)



# Autocorrelation

- Assess for autocorrelation using
  - Cumby-Huizinga test for autocorrelation
    - **STATA command:** `actest , lag(#)`
  - Visual assessments
    - Correlogram
      - **STATA command:** `corrgram varname`
    - Autocorrelation (ACF) plots
      - **STATA command:** `ac varname`
    - Partial autocorrelation (partial ACF) plots
      - **STATA command:** `pac varname`



# Cumby-Huizinga test for autocorrelation

**STATA command:** `actest, lags(10)`

Cumby-Huizinga test for autocorrelation

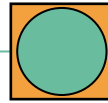
H0: variable is MA process up to order q

HA: serial correlation present at specified lags >q

H0: q=0 (serially uncorrelated) HA: s.c. present at range specified				H0: q=specified lag-1 HA: s.c. present at lag specified			
lags	chi2	df	p-val	lag	chi2	df	p-val
1 - 1	15.642	1	0.0001	1	15.642	1	0.0001
1 - 2	15.937	2	0.0003	2	4.540	1	0.0331
1 - 3	17.735	3	0.0005	3	4.382	1	0.0363
1 - 4	17.751	4	0.0014	4	2.366	1	0.1240
1 - 5	18.666	5	0.0022	5	2.678	1	0.1018
1 - 6	19.174	6	0.0039	6	0.969	1	0.3249
1 - 7	19.661	7	0.0064	7	0.023	1	0.8802
1 - 8	19.661	8	0.0117	8	0.003	1	0.9588
1 - 9	21.150	9	0.0120	9	0.943	1	0.3316
1 - 10	22.100	10	0.0146	10	0.522	1	0.4701

Test allows predetermined regressors/instruments

Test requires conditional homoskedasticity



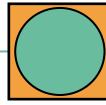
# Final Model

**STATA command:** itsa conANC4, single treatid(20) trperiod(11) lag(3) figure

Regression with Newey-West standard errors				Number of obs	=	44
maximum lag: 3				F( 3, 40)	=	56.98
				Prob > F	=	0.0000
conANC4	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]	
_t	.0000329	7.95e-06	4.14	0.000	.0000168	.0000489
_x11	.0031145	.0008804	3.54	0.001	.0013351	.0048939
_x_t11	.0001136	.0000504	2.25	0.030	.0000118	.0002154
_cons	.0013329	.0000433	30.79	0.000	.0012454	.0014204

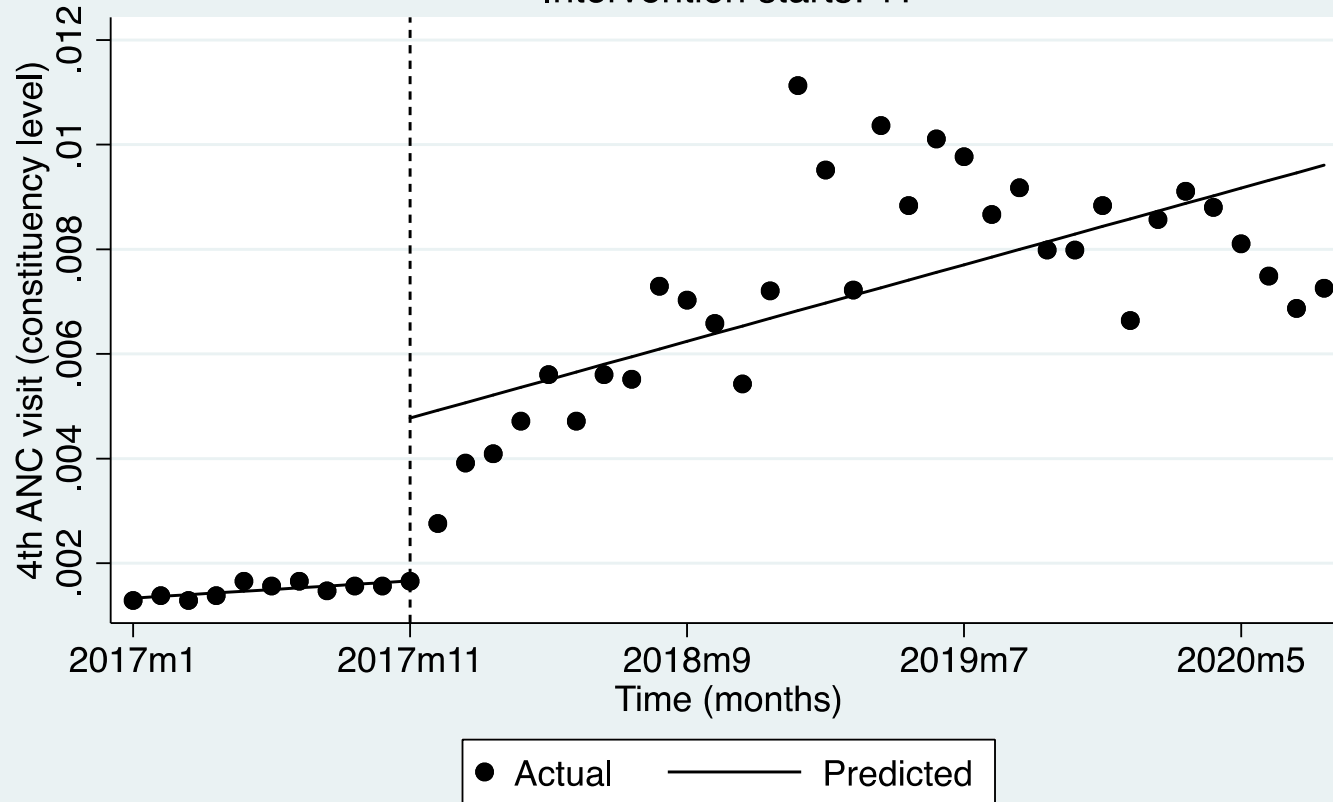
Variable	Explanation	Coefficient	P> t	[95% Conf. Interval]
_t	Existing trend ( $\beta_1$ )	0.0000329	0.000	[0.0000168, 0.0000489]
_x11	Level change ( $\beta_2$ )	0.0031145	0.001	[0.0013351, 0.0048939]
_x_t_11	Trend change ( $\beta_3$ )	0.0001136	0.030	[0.0000118, 0.0002154]
_cons	Existing level ( $\beta_0$ )	0.0013329	0.000	[0.0012454, 0.0014204]





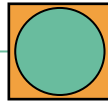
## Gold-C. IX: ANC (4th) visits over time

Intervention starts: 11



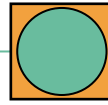
Regression with Newey-West standard errors - lag(3)

**Interpretation:** After the policy change, there was a sustained increase in the constituency rate of the 4<sup>th</sup> ANC visit, by about 0.3%. There was an increase in the trend by 0.01% per month.



# Predicted changes

- Impact of policy 1 year later
  - *Absolute change = fitted – counterfactual*
  - *Relative change =  $\frac{\text{fitted} - \text{counterfactual}}{\text{counterfactual}}$*



## Predicted changes

$$y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t$$

- 1 year (12 months) after policy
  - time since study start = 11 + 12 = 23
  - time since policy start = 12

- Predict counterfactual at 1 year

- $X = 0, T = 23, T \text{ in } XT = 12$

$$y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t$$

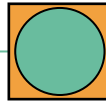
$$y_t = \beta_0 + (\beta_1 \times 23) + (\beta_2 \times 0) + (\beta_3 \times 0 \times 12)$$

$$y_t = \beta_0 + (\beta_1 \times 23)$$

```
. lincom _b[_cons]+(_b[_t]*23)
```

```
( 1) 23*_t + _cons = 0
```

conANC4	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	.0020894	.000154	13.57	0.000	.0017781	.0024006



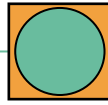
# Predicted changes

- Predict fitted value at 1 year
  - $X = 1$ ,  $T = 23$ ,  $T$  in  $XT = 12$

$$\begin{aligned}y_t &= \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t \\&= \beta_0 + (\beta_1 \times 23) + (\beta_2 \times 1) + (\beta_3 \times 1 \times 12) \\&= \beta_0 + (\beta_1 \times 23) + \beta_2 + (\beta_3 \times 12)\end{aligned}$$

```
.  
. lincom _b[_cons]+(_b[_t]*23)+_b[_x11]+(_b[_x_t11]*12)  
  
( 1) 23*_t + _x11 + 12*_x_t11 + _cons = 0
```

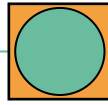
conANC4	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	.0065666	.0005029	13.06	0.000	.0055502	.007583



## Predicted changes

- Impact of policy 1 year later
  - *Absolute change* =  $0.0065666 - 0.0020894$   
= .0044772
  - *Relative change* =  $\frac{0.0065666 - 0.0020894}{0.0020894}$   
= 2.1428161

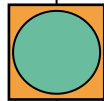
**Interpretation:** In the 1<sup>st</sup> year after the policy, the rate of 4<sup>th</sup> ANC visits was 0.44% more than would have been expected if the policy was not implemented. This represents a 214% increase.



# Summary

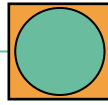
- In single group ITS, there is one cohort that experienced the event/intervention/policy.
- ITSA command in STATA can be used for single and multiple group comparisons
  - Controls for autocorrelated disturbances and estimate the treatment effects over multiple periods.
  - Generates graphics of the ITS
- Post estimation commands such as the Cumby-Huizinga test for autocorrelation can be used to determine the maximum number of lags to include in the analysis
- In addition to determining the impact, we can predict the absolute and relative change

# Resources



- Background information sheet
  - Overview of time series and interrupted time series
  - List of useful resources
  - Summary of studies using interrupted time series analysis
- PowerPoint presentation
- Sample dataset\*
- Sample do file\*

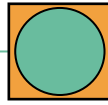
\* Resources are available for Part 2 and 3 of the ITS video tutorial only



# Looking ahead

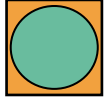
- Part 1:
  - Interrupted time series overview
  - Important considerations
  - Preparing the dataset
- Part 2:
  - Interrupted time series analysis: Single group
- Part 3:
  - Interrupted time series analysis: Multiple group





# References

- Baum, C.F., Schaffer, M.E. 2013. `actest`: Stata module to perform Cumby-Huizinga general test for autocorrelation in time series. <http://ideas.repec.org/c/boc/bocode/s457668.html>
- 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>
- Cochrane Effective Practice and Organisation of Care (EPOC). (2017). Interrupted time series (ITS) analyses: EPOC Resources for review authors. <https://www.coursehero.com/file/74328642/interrupted-time-series-analysesdocx/>
- Law, M. Policy Analysis Using Interrupted Time Series. University of British Columbia. <https://www.edx.org/course/policy-analysis-using-interrupted-time-series>
- Linden, A. (2015). Conducting interrupted time-series analysis for single- and multiple-group comparisons. *The Stata Journal*, 15(2), 480-500. <https://www.stata-journal.com/article.html?article=st0389>
- Penfold, R. B., & Zhang, F. (2013). Use of interrupted time series analysis in evaluating health care quality improvements. *Academic pediatrics*, 13(6 Suppl), S38–S44. <https://doi.org/10.1016/j.acap.2013.08.002>



This presentation was produced with the support of the United States Agency for International Development (USAID) under the terms of the Data for Impact (D4I) associate award 7200AA18LA00008, which is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with Palladium International, LLC; ICF Macro, Inc.; John Snow, Inc.; and Tulane University. The views expressed in this publication do not necessarily reflect the views of USAID or the United States government.

[www.data4impactproject.org](http://www.data4impactproject.org)

