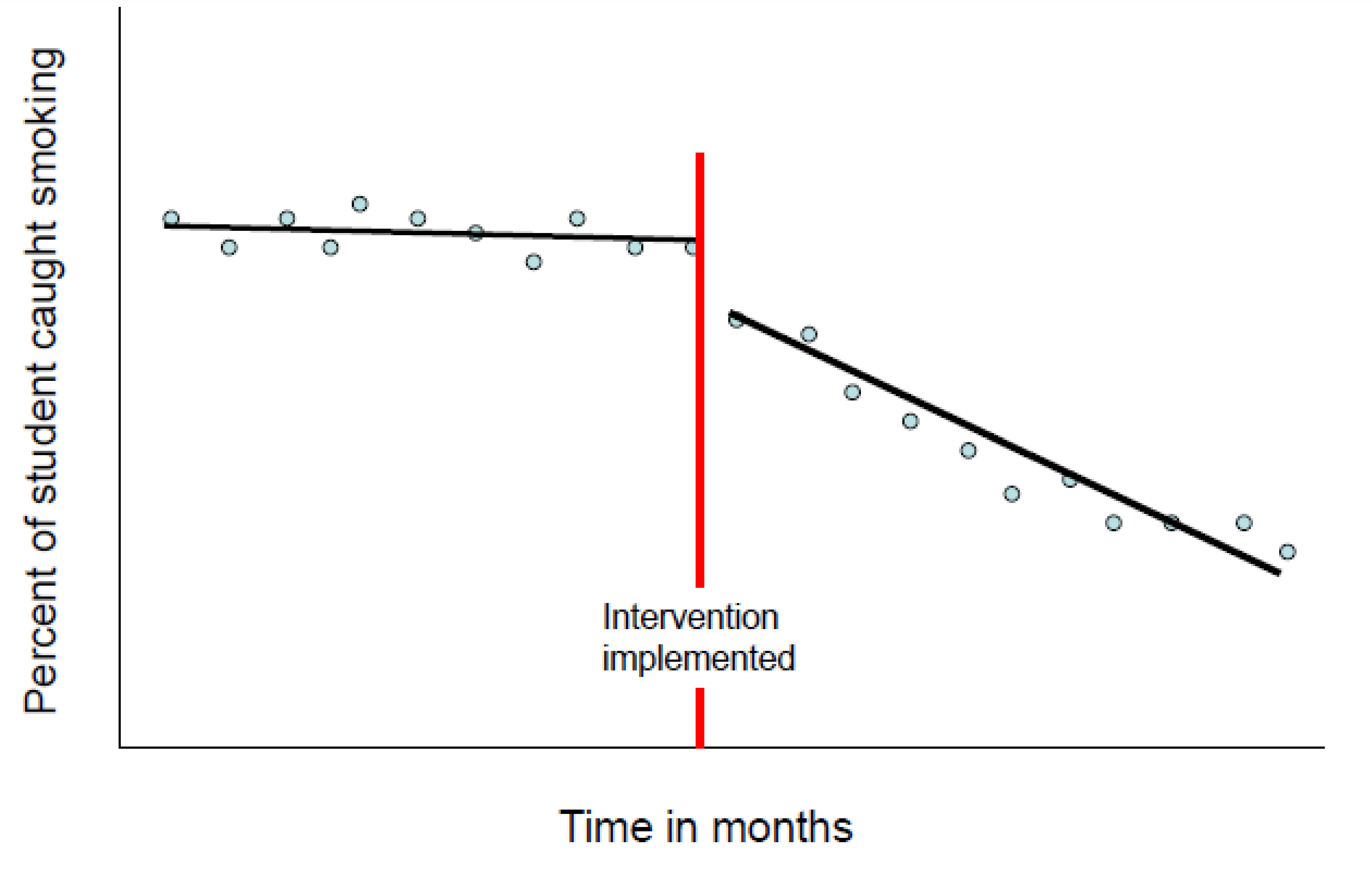
The Strategy Unit.

2.02 The interrupted time series design: statistical analysis

NHS-R workshop no. 4 | February 2019

NHS
Midlands and Lancashire
Commissioning Support Unit


Percent of student caught smoking
Intervention
implemented
Time in months

Statistical analysis

Two parameters define each segment of a time series: level and trend (or slope).

It's possible to make an assessment of changes in level and/or trend by visual inspection of the time-series. But unable to state whether changes in level and trend are the result of (1) chance, or (2) factors other than the intervention.

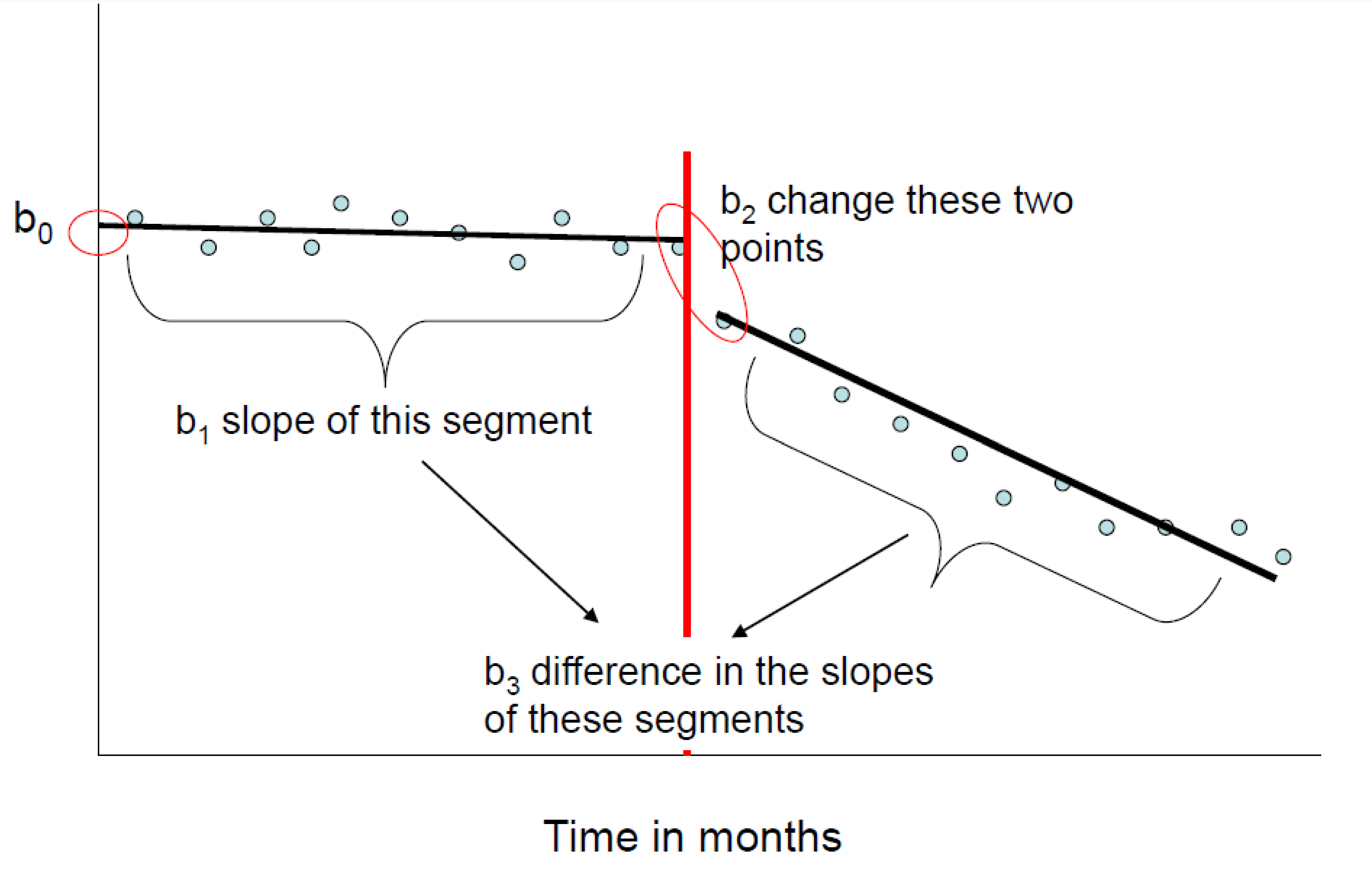
The most common method for analysis of ITS is "segmented" regression analysis. Linear regression models are used to estimate level and trend in the pre-intervention segment and changes in level and trend after the intervention.

Such analyses can be used to assess chance and control for other effects—segmented regression controls for baseline level and trend (secular changes that may have occurred in the absence of the intervention).

Segmented regression

In its simplest form, an ITS is modelled using a “segmented” regression model that includes only three time based covariates

1. the pre-intervention slope quantifies the trend for the outcome before the intervention
2. the level change is an estimate of the change in level that can be attributed to the intervention, between the time points immediately before and immediately after the intervention, and accounting for the pre-intervention trend
3. the change in slope quantifies the difference between the pre-intervention and post-intervention slopes



b0
b2 change these two points
b1 slope of this segment
b3 difference in the slopes of these segments
Time in months

Impact: change in level

b_1 = baseline slope

$$Y_t = b_0 + b_1 T + b_2 D + e_t$$

b_2 = change in level

Where

T is time from the start of the observation period

D is a dummy variable (coded 0, 1) denoting pre or post intervention

Example data frame

Year	Month	Time elapsed (T)	Intervention (D)	Outcome (Y)
…				
2015	6	30	0	914
2015	7	31	0	808
2015	8	32	0	937
2015	9	33	0	840
2015	10	34	0	916
2015	11	35	0	828
2015	12	36	0	845
2016	1	37	1	860
2016	2	38	1	796
2016	3	39	1	825
2016	4	40	1	841
2016	5	41	1	765
2016	6	42	1	782

Impact: change in level & change in slope

(b_1+b_3) = new slope

b_1 = baseline slope

b_3 = change in slope

$$Y_t = b_0 + b_1 T + b_2 D + b_3 P + e_t$$

b_2 = change in level

Where

T is time from the start of the observation period

D is a dummy variable (coded 0, 1) denoting pre or post intervention

P is time since the intervention (coded 0 prior to the intervention)

Example data frame

Year	Month	Time elapsed (T)	Intervention (D)	Time since intervention (P)	Outcome (Y)
…					
2015	6	30	0	0	914
2015	7	31	0	0	808
2015	8	32	0	0	937
2015	9	33	0	0	840
2015	10	34	0	0	916
2015	11	35	0	0	828
2015	12	36	0	0	845
2016	1	37	1	37	860
2016	2	38	1	38	796
2016	3	39	1	39	825
2016	4	40	1	40	841
2016	5	41	1	41	765
2016	6	42	1	42	782

Reporting effects

Can compare observed post-intervention values with post-intervention values estimated from the level and trend of the pre-intervention period i.e. the counterfactual—what would the trend have looked like in the absence of the intervention.

Identified effect can be expressed as:

1. Absolute difference between estimated outcome based on the intervention and the counterfactual value

2. Ratio of the estimated and counterfactual values

Need to select a point in time e.g. 6 months post-intervention, 12 months post-intervention.

Technical considerations

Power—is greatest when the intervention occurs in the middle of the series. A stable baseline increases power to detect change—basing each observation on a larger sample size will reduce variability, therefore consider aggregating data.

Linearity—standard regression models assume a linear trend in the outcome within each segment. The assumption of linearity often may hold only over short intervals.

Autocorrelation—OLS regression assumes that the error terms associated with each observation (time point) are uncorrelated. This assumption may not hold for time series data.

Seasonality—time of year may influence the outcome variable. Detection requires observations spanning multiple seasonal periods. If present should be controlled for when estimating effects.