Introduction to Interrupted Time Series

Part I. Concepts

Joint CAPS/TAPS Methodology Seminar

January 19, 2016

Steve Gregorich
Main Goals

• Brief(ish) introduction to interrupted time series (ITS) designs

• Comparing ITS to some other quasi-experimental designs that do not include control groups.
Outline

• First look at an Interrupted Time Series (ITS)

• Gold standard: The randomized controlled trial design

• Generic threats to internal validity: RCTs

• Some quasi-experimental designs (QED) with no control group

• Generic threats to internal validity: QEDs

• The ITS design; real examples of some archetypal outcomes

• Bolstering the ITS design

• ITS analysis, very briefly

• Summary
First look at an Interrupted Time Series Design


April 2005: FDA issued an advisory and black box warning
Risks of ↑ mortality: atypical anti-psychotic use: elderly patients w/ dementia

The impact of these warnings on atypical drug use was unknown
Gold Standard: The Randomized Controlled Trial Design

Rnd \begin{align*}
\text{Intv:} & \quad O_{t_1} \quad Tx \quad O_{t_2} \\
\text{Ctrl:} & \quad O_{t_1} \quad O_{t_2}
\end{align*}

- Rnd: Equivalent groups at $t_1$.

- If 'closed-system' maintained,
  then solid basis for causal inference about Tx effects

I.e., internal validity
Generic Threats to Internal Validity

Focal (for today)

• **Selection**: participant characteristics systematically differ across groups

• **History**: events acting upon population & co-occurring with Tx

• **Maturation**: natural changes in sampled Pts across time

• **Testing**: repeated exposure to a test may affect assessment
Generic Threats to Internal Validity

*Others*—almost universally problematic

- **Instrumentation**: the nature of a measure changes across time, such that the validity of repeated assessments may be questioned

- **Ambiguous temporal sequencing of variables**: $X \rightarrow Y$, or $Y \rightarrow X$?

- **Regression**: Pts with initial extreme values may 'regress'

- **Attrition**: if systematically correlated with Tx or outcomes

*All threats (Focal and Others) can combine additively or interactively*
RCT and 'Focal' Threats to Internal Validity

\[
\begin{align*}
\text{Intv:} & \quad O_{t1} \quad \textbf{Tx} \quad O_{t2} \\
\text{Rnd} & \quad \begin{cases} \\
\text{Ctrl:} & \quad O_{t1} \quad O_{t2} \\
\end{cases}
\end{align*}
\]

**Selection:** randomization should address

**History:** synchronized assessments should address

**Maturation:** randomization & synchronized assessments should address

**Testing:** parallel assessment schedule should address
RCT and 'Other' Threats to Internal Validity

\[
\begin{align*}
\text{Rnd} & : O_{t1} & \text{Tx} & : O_{t2} \\
\text{Ctrl} & : O_{t1} & O_{t2}
\end{align*}
\]

**Instrumentation**: addressed, as long as measures are relevant to targeted constructs

**Ambiguous temporal sequencing**: longitudinal design addresses

**Regression**: randomization and parallel assessments should address, even if extreme groups are targeted for recruitment

**Attrition**: always a concern; to be dealt with in a principled fashion
Some Longitudinal QED designs w/ no Control Group
Often, QI study designs do not employ a control group

<table>
<thead>
<tr>
<th></th>
<th>One Sample, Longitudinal</th>
<th>Multiple-Cross Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretest-posttest</td>
<td>$O_{t1}$ Tx $O_{t2}$</td>
<td>$O_{t1}$ Tx $O_{t2}$</td>
</tr>
<tr>
<td>pre-post w/ multi-pre</td>
<td>$O_{t0}$ $O_{t1}$ Tx $O_{t2}$</td>
<td>$O_{t0}$ $O_{t1}$ Tx $O_{t2}$</td>
</tr>
<tr>
<td>repeated Tx</td>
<td>$O_{t1}$ Tx $O_{t2}$ Tx $O_{t3}$ Tx $O_{t4}$</td>
<td>$O_{t1}$ Tx $O_{t2}$ Tx $O_{t3}$ Tx $O_{t4}$</td>
</tr>
</tbody>
</table>

• Many other designs exist
Summary: Internal Validity Threats w/ no Control Group

One-Sample, Longitudinal QEDs

<table>
<thead>
<tr>
<th></th>
<th>selection</th>
<th>history</th>
<th>maturation</th>
<th>testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_{t1}$ $Tx$ $O_{t2}$</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>$O_{t0}$ $O_{t1}$ $Tx$ $O_{t2}$</td>
<td>x</td>
<td>reduced</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>$O_{t1}$ $O_{t2}$ $Tx$ $O_{t3}$ $Tx$ $O_{t4}$</td>
<td>greatly reduced</td>
<td></td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

Multiple-Cross Sectional, QEDs

<table>
<thead>
<tr>
<th></th>
<th>selection</th>
<th>history</th>
<th>maturation</th>
<th>testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_{t1}$ $Tx$ $O_{t2}$</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_{t0}$ $O_{t1}$ $Tx$ $O_{t2}$</td>
<td>x</td>
<td>x</td>
<td>reduced</td>
<td></td>
</tr>
<tr>
<td>$O_{t1}$ $O_{t2}$ $Tx$ $O_{t3}$ $O_{t4}$</td>
<td>x</td>
<td>reduced</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Interrupted Time Series Design

• Longitudinal

\[ O_{t1} \quad O_{t2} \quad O_{t3} \quad O_{t4} \quad O_{t5} \quad T_x \quad O_{t6} \quad O_{t7} \quad O_{t8} \quad O_{t9} \quad O_{t10} \]

• Multiple cross-section

\[ O_{t1} \quad O_{t2} \quad O_{t3} \quad O_{t4} \quad O_{t5} \quad T_x \quad O_{t6} \quad O_{t7} \quad O_{t8} \quad O_{t9} \quad O_{t10} \]

Either way, it can be a strong design
ITS Example 1: Charging for directory assistance (DA)

- A change in level at intervention onset (March 1974). Y-axis: # calls

**FIGURE 6.1** The effects of charging for directory assistance in Cincinnati.
ITS Example 1: Charging for directory assistance (DA)

- Immediate large drop in number of calls, March 1974

  **Selection** implausible:
  pre and post samples likely the same

  **Attrition** implausible
  New charges unlikely to prompt phone disconnections

  **Maturation** implausible
  no known maturation process could account for drop in calls

  **History** implausible
  unless another hypothetically causal event can be identified

  **Testing** implausible
  E.g., if phone co. changed salience of DA charges on phone bills

  **Regression to the mean** implausible:
  pre- trend suggested high call rates for many years
ITS Example 2: New Law Re. Sexual Assault Reporting

- Change in slope at intervention onset. Y-axis: # reported sexual assaults

Canada 1983 included highly publicized provisions to increase reporting to police

Note seasonal variation
ITS Example 2: New Law Re. Sexual Assault Reporting

- Immediate change in slope from flat to positive, 1983

  **Maturation** implausible
  no known maturation process could account for change in slope

  **History** implausible
  unless another hypothetically causal event can be identified

  **Instrumentation** possible.
  The new law changed the categories of reportable sexual assault
  1. wives could charge husbands with sexual assault
  2. included assaults against both males and females

  Authors showed that, in the post-intervention period, *suspects* who were women or husbands did not increase sufficiently to explain the pattern of results.
ITS Example 3: Alcohol warning label re. prenatal drinking

- Weak, Delayed, Ambiguous Effects. Y-axis: Prenatal Drinking Score

![Graph showing trends in prenatal drinking over time.](image)

**FIGURE 6.3** The effects of an alcohol warning label on prenatal drinking


DV: Alcohol consumption 2 weeks prior to 1st prenatal visit
ITS Example 3: Alcohol warning label re. prenatal drinking

Delayed effect, authors argued, 7 months after implementation
Law affected new containers, not those already on store shelves
When asked, women were not aware of the labels until +4 months

Maturation threat? Drinking was decreasing prior to the intervention
Analysis: delayed Tx slope was stronger than pre-Tx slope

Seasonal variation threat?
Typically, drinking increased during Nov/Dec and Summer
Given intervention timing & results, seasonality not a strong threat
November onset: Seasonal & treatment effects in opposite directions
If intervention implemented in Feb or Sept, then seasonal effects might be misinterpreted as intervention effects.
Much lower rates in Nov/Dec ‘90 and Summer ‘91
Compared to holiday/summer periods in previous years
ITS Example 4: Pay-for-performance & BP control

• No effect observed
ITS advantages over pre-test / post-test design: Simplified

Scenario #1: intervention effect observed: immediate change in slope

. ITS would identify the intervention effect

. A simple pre-post test design would not.

Comparing the pre- and post- means (black dots) suggests no overall pre-post difference
ITS advantages over pre-test / post-test design: Simplified

Scenario #2: no intervention effect

. ITS would identify the lack of intervention effect

A simple pre-post test design would suggest an intervention effect. Comparing pre- and post- means (black dots) suggests a post-test increase in outcome level
Summary, So Far

• ITS design can provide a good basis for drawing causal inferences if...
  . observed changes are well timed with intervention onset
  . alternative explanations (threats to internal validity) are *implausible*

• However, even under those circumstances threats to internal validity may still operate, e.g., the seemingly implausible may obtain

• Next: ways to bolster the ITS design
Bolstering the ITS Design

• Non-equivalent no-treatment control group
• Non-equivalent dependent variables
• Removing a treatment at a known time
• Multiple replications
• Switching replications
Bolstering the ITS Design

Non-equivalent, no-treatment control group

\[
\begin{array}{cccccccc}
  O_{t1} & O_{t2} & O_{t3} & O_{t4} & O_{t5} & Tx & O_{t6} & O_{t7} & O_{t8} & O_{t9} & O_{t10} \\
  O_{t1} & O_{t2} & O_{t3} & O_{t4} & O_{t5} & Tx & O_{t6} & O_{t7} & O_{t8} & O_{t9} & O_{t10} \\
\end{array}
\]

Concept
• I.e., add a group hypothetically unaffected by the intervention

Example: L. Karliner (PI)
  Impact of hospital "bedside interpreter" on LEP patient outcomes
  (add a non-equivalent no-treatment control group of EP patients)

Most notably, this helps to diagnose **history** threats (made-up examples)
Bolstering the ITS Design

*Non-equivalent, no-treatment control group*

Addressing threats to internal validity

**History**

EP & LEP patients live in same city, treated in same hospital
So, many potential historical effects would be equivalent

**Instrumentation**

If hospital charts and billing record systems changed, then the change should affect both EP and LEP groups

**Selection:** here non-equivalence of groups is intentional
A problem could arise if history and selection effects interacted to produce differential group effects around the time of intervention onset
Bolstering ITS Design:

**Non-equivalent dependent variables**

\[
\begin{array}{cccccccccc}
O_{At1} & O_{At2} & O_{At3} & O_{At4} & O_{At5} & Tx & O_{At6} & O_{At7} & O_{At8} & O_{At9} & O_{At10} \\
O_{Bt1} & O_{Bt2} & O_{Bt3} & O_{Bt4} & O_{Bt5} & Tx & O_{Bt6} & O_{Bt7} & O_{Bt8} & O_{Bt9} & O_{Bt10}
\end{array}
\]

**Concepts**
Add an outcome hypothesized to be unaffected by the intervention, Main & non-equiv. DVs should be equally subject to validity threats
Main & non-equivalent DVs should be conceptually related

**Example: Directory assistance (DA)**
The DA charge was only for local numbers, not long distance (LD)
The author noted that only local DA calls changed, not LD DA calls
Bolstering ITS Design: British Breathalyzer example

*Non-equivalent dependent variables*

**FIGURE 6.6** The effects of the British Breathalyzer crackdown on traffic casualties during weekend nights when pubs are open, compared with times when pubs were closed.


Closed hours: e.g., commute times

History threats (e.g., ↑ speed traps, safer cars) should affect all serious accidents regardless of time of day.
Bolstering the ITS Design

Removing a treatment at a known time

\[ O_{t1} \quad O_{t2} \quad O_{t3} \quad O_{t4} \quad Tx \quad O_{t5} \quad O_{t6} \quad O_{t7} \quad Tx \quad O_{t8} \quad O_{t9} \quad O_{t10} \]

**FIGURE 6.8** The effects of psychiatric crisis intervention on hospitalization

Bolstering the ITS Design

*Removing a treatment at a known time*

Addressing threats to internal validity

**History**
If the hypothesized response pattern obtained, then a credible historical threat would require one or more effects that
i. operate in different directions at different times, and
ii. are well-timed with intervention onset and removal

**Selection**
If this treat were credible, then it would require different types of people to enter and leave the population at different times

**Instrumentation**
A less plausible threat with this design

Non-equivalent control group would provide additional strength regarding possible History and Maturation (e.g., seasonal trend) effects
Bolstering the ITS Design

Adding multiple replications

<table>
<thead>
<tr>
<th>$O_{t1}$</th>
<th>$O_{t2}$</th>
<th>$T_x$</th>
<th>$O_{t3}$</th>
<th>$O_{t4}$</th>
<th>$T_x$</th>
<th>$O_{t5}$</th>
<th>$O_{t6}$</th>
<th>$T_x$</th>
<th>$O_{t7}$</th>
<th>$O_{t8}$</th>
<th>$T_x$</th>
<th>$O_{t9}$</th>
<th>$O_{t10}$</th>
</tr>
</thead>
</table>

**FIGURE 6.9** The effects of treatment for inflammation of continent ileostomy. In the graphs, the letter $T$ indicates the time period during which treatment occurred.

Bolstering the ITS Design

Adding switching replications

2 or more nonequivalent groups w/ staggered intervention introduction

\[
\begin{array}{cccccccc}
O_{t1} & O_{t2} & O_{t3} & Tx & O_{t4} & O_{t5} & O_{t6} & O_{t7} & O_{t8} & O_{t9} & O_{t10} \\
O_{t1} & O_{t2} & O_{t3} & O_{t4} & O_{t5} & O_{t6} & O_{t7} & Tx & O_{t8} & O_{t9} & O_{t10}
\end{array}
\]

**FIGURE 6.10** The effects of the introduction of television on property crime rates in cities in which television was introduced in 1951 versus 1955

Bolstering the ITS Design

*Longitudinal cohort studies*

Patient-level data, longitudinally collected before and after an intervention (e.g., surgery).

E.g., prospective cohort study where patients are followed for 10 years. Some have a surgical intervention during the study observation period. Interest is in comparing pre- and post-surgical outcome trends.

Possible to model patient-level time series as switching replications. However, individual patient time series may be very noisy and some patients will have very few pre- or post-surgery observations.

An alternative is to code the time of each patient’s surgery as time=0, pool data across patients, and model using repeated measures regression.
Analysis of data from ITS designs

Originally, time-series analysis, a modeling framework from econometrics, was used almost exclusively.

Analysis Alternatives
  . Repeated measures models
  . Segmented linear regression

We will discuss some analysis options during the 2nd hour.
Summary

• ITS vs. other QED wrt threats to internal validity
  . ITS far superior to pre-/post-test type designs with no control group
  . ITS better than pre-post designs with an unmatched, non-randomized control group
  . ITS can be better than pre-post designs with a matched non-randomized control group sample

• A suggested 'minimal' ITS design
  . intervention onset at a single point in time
  . intervention delivered to one population
  . add a non-equivalent control group
  . add non-equivalent outcomes, if possible

• Often attainable advanced design element: Switching replications
  . A natural addition when working in multiple practices within a system, multiple hospital systems, etc.
Summary

• Units of analysis
  Outcomes often aggregated monthly, quarterly, or annual summaries
e.g., annual incidence of a specific condition,
total quarterly costs, average (or median)

• Trade-off between length of observation, level of aggregation, noise

• Signal-to-noise ratio
  Series of patient-level outcomes can be very noisy, even in aggregate
  This may require large numbers of observations
  May not be feasible for some studies with primary data collection
  Longitudinal (vs. multiple cross-sectional) data help wrt power
  Experience is the best guide

• Medical records, billing data, claims data, administrative data
  Opportunities to evaluate clinical policy changes, either
    . Truly retrospectively
    . Semi-prospectively, with the aid of retrospective pre- data

END