Causal Learning With Interrupted Time Series

Yiwen Zhang Benjamin M. Rottman, Ph.D.

University of Pittsburgh

INTRODUCTION

Interrupted time series

- Interrupted time series analysis (ITSA) is a statistical procedure that evaluates whether an intervention causes a change in the level and/or slope of the time series. A simple ITSA is modeled with three components: the time since the start of first observation, level change and slope change[1].
- Previous research found that people failed to control for pre-intervention slope and only compared the mean level of outcomes in pre- and post- intervention period[2,3]. However, the research didn't investigate conditions with level changes or slope changes.
- Can people learn true causation from various interrupted time series scenarios?

Presentation format of time series data

- Dynamic presentations helped people accurately learn causal relationships by focusing on changes in the cause and effect, whereas static and numerical presentations led them to focus on the simple correlation and not account for trends[4].
- Will presentation formats affect causal learning with interrupted time series?

Three potential theories of learning interrupted time series:

- Formal interrupted time series analysis (ITSA): looking for changes in the intercept or slope after the intervention compared to before.
- After-minus-before heuristic[2]: comparing the mean of the data after the intervention vs. before.
- Post-intervention trend: simply focusing on the slope of the post-intervention trend.

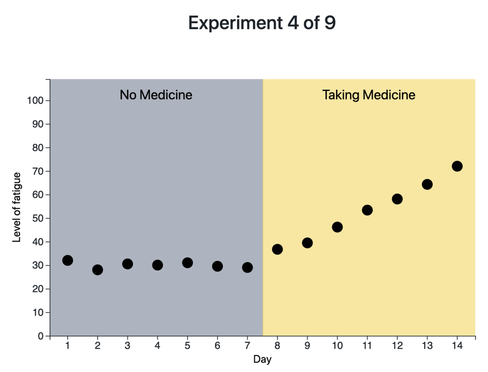
	Model Prediction			Empirical
Condition	ITSA	After-before	Post. Trend	Data
A. Flat	0	0	0	0
B. Pre-intervention Slope	0	+	+	+
C. Intercept Change	+	+	0	+
D. Slope Change	+	+	+	+
E. Slope Change (Maintain)	+	-	0	0
F. Slope Change (C. - PS)	+	+	+	+
G. Intercept Change (C. - PS)	+	+	+	+
H. Intercept Change (I. - PS)	+	-/0	-	-/0
I. Slope Change (I. - PS)	+	-/0	+	+

METHODS

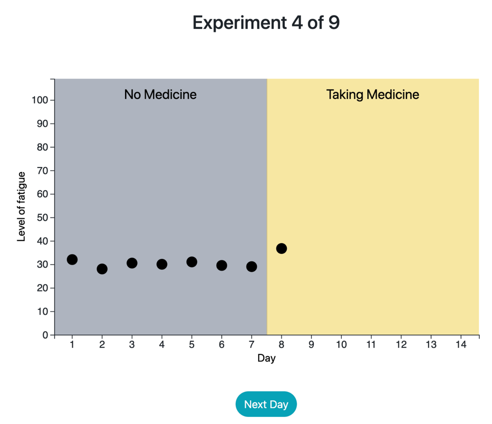
Procedure

402 participants from Mechanical Turk completed the study. They were told to imagine that they work for a medicine company that is testing the efficacy of new medications. They reviewed 9 datasets in randomized order, each with a different medicine (e.g., SNP27), and a different symptom (e.g. headache, back pain). Each depicted the data for a single patient over an initial week without and one week with the medicine.

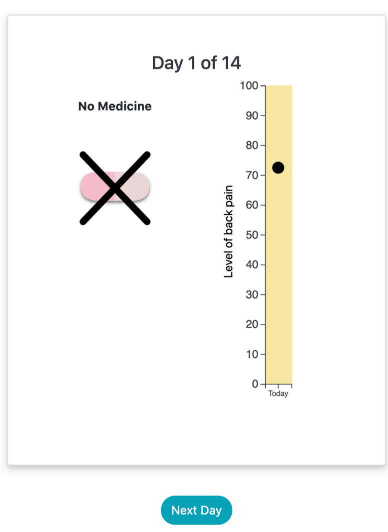
Static Graph



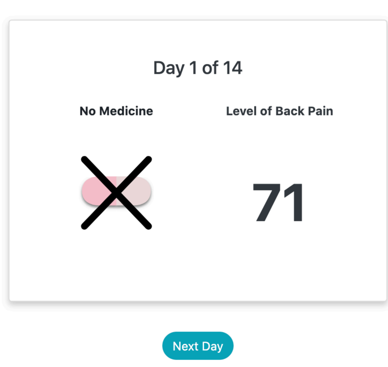
Dynamic Graph



Trial by Trial - Dot



Trial by Trial - Number



Design 4*9 Mixed Design.

Within subject manipulation: 9 time series conditions (See Figure on the right). The theory predictions of each condition is shown in the table.

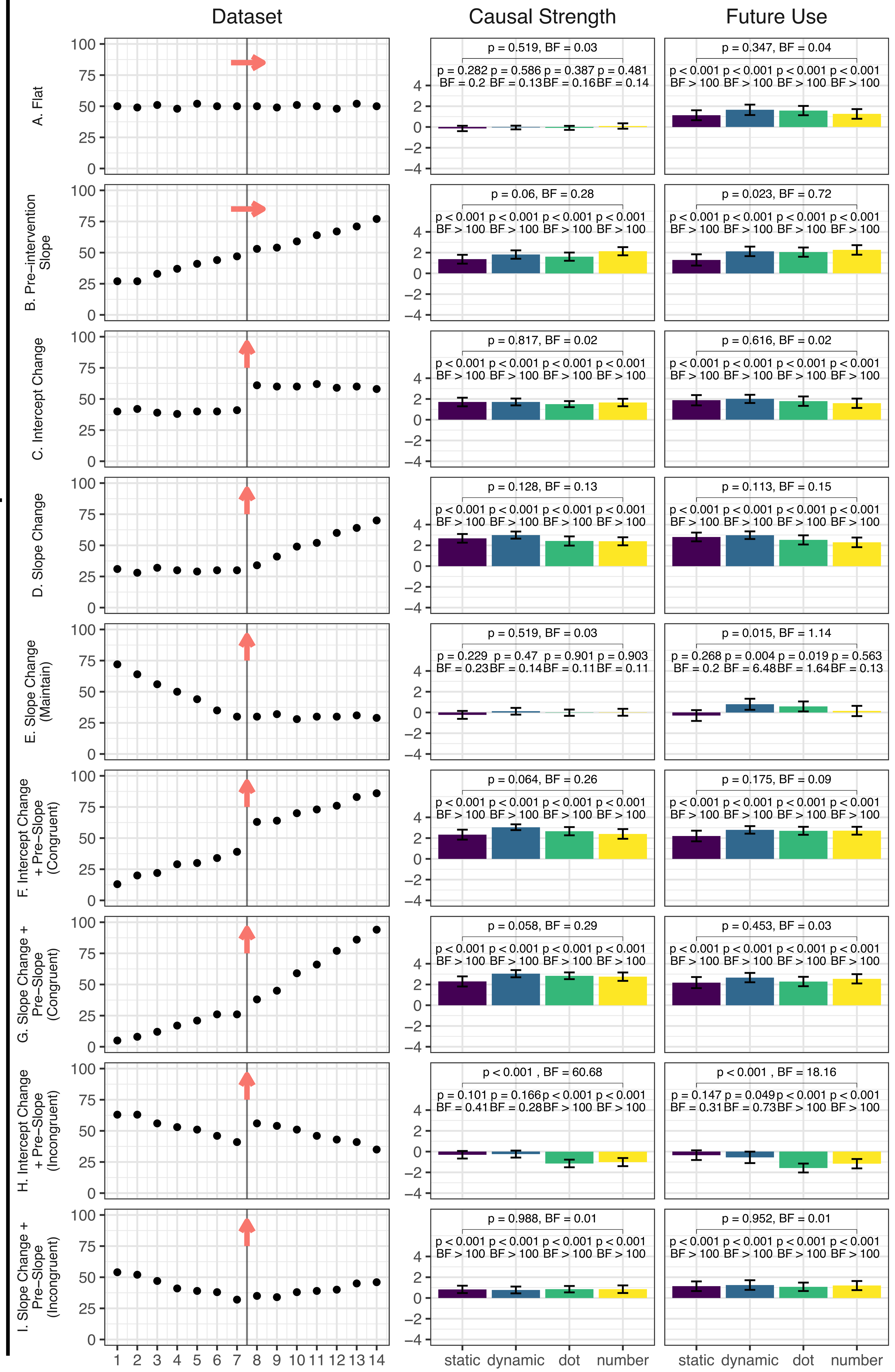
Between subject manipulation: 4 presentation formats (See Figure on the bottom)

- **Static Graph**: All 14 observations were presented at once in a dot chart.
- **Dynamic Graph**: Identical to the static graph condition, except that a data point was added to the graph each time participants clicked a button.
- **Trial by Trial – Dot (TbT-Dot)**: Participants saw one observation per trial, an icon indicating medicine (or no medicine) and a narrow bar chart indicating the level of outcome, and clicked a button to see the next trial.
- **Trial by Trial – Number (TbT-Number)**: Identical to the Dot condition, except that the dot chart was replaced by a number of the level of the outcome.

Dependent Measures

- **Causal Strength**: Participants answered "Did taking [medicine] cause the [symptom] to get better or worse?" on a scale from 1-9 scale: 1 (much worse - higher), 5 (no influence), to 9 (much better - lower).
- **Future Use Strength**: Participants answered "Do you think this patient should continue to take the medicine to treat the symptom?" on a 1-9 scale: 1 (definitely stop), 5 (unsure), to 9 (definitely).

For each time series condition and presentation formats, we tested if participants judged the intervention as effective by conducting t-test against 0. The t-test results are on the top of each bar. Then we tested the effect of presentation formats by conducting ANOVA for each time series condition and each measure. The red arrow in the Dataset column indicate the formal judgments of the intervention. ↑ means positive causation and → means no causation.



RESULTS

For each interrupted time series condition, we included parallel datasets simply involved flipping the Y axis, for generality. We centered the response and inversely coded the responses for negative conditions.

Nine Datasets

- An ITSA approach accounts for the results 6 out of 9 conditions and failed to explain the results in Condition B, E and H.
- The after-minus-before theory agrees with ITSA and also correctly predicted results in 5 conditions (A, C, D, F, G). Unlike ITSA, it also correctly predicted the results in Condition B.
- The post-intervention-trend theory also agrees with ITSA in 5 out of the 9 conditions (A, D, F, G, and I). It correctly predicted the results for 7 conditions. However, this theory cannot explain the results in Condition C.

Presentation Formats

- For eight out of the nine conditions there were no reliable effects of presentation formats.
- We did find a main effect of presentation format in Condition H. The judgements with the static and dynamic graph formats were close to zero but the TbT-dot and TbT-number formats were less than zero.

DISCUSSION

This is the first study to systematically investigate causal learning under interrupted time series data.

Comparison of Models

- None of these three theories can explain all the results, which means that either participants used a combination of these theories, that there are mixtures of different groups of participants, or that there are other theories that better explain the results.

Effects of Incongruency

- Incongruence is when the the influence of the intervention opposes the direction of the pre-intervention slope.
- Participants had difficulty correctly assessing causality in the three `incongruent' conditions (E, H, I).

Effect of presentation formats

- The format only affected learning in one of the datasets.

REFERENCES

[1]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.

[2]White, P. A. (2015). Causal judgements about temporal sequences of events in single individuals. *Quarterly Journal of Experimental Psychology, 68(11)*, 2149-2174.

[3]White, P. A. (2017). Causal judgments about empirical information in an interrupted time series design. *Quarterly Journal of Experimental Psychology, 70*(1), 18-35.

[4]Soo, K. W., & Rottman, B. M. (2020). Distinguishing causation and correlation: Causal learning from time-series graphs with trends. *Cognition*, *195*, 104079.