HOW WE IDENTIFY CAUSE-EFFECT RELATIONSHIPS GIVEN EVIDENCE?
Extended Abstract

Motivation

The ability to understand causality is essential for accurate predictions and informed decision-making. While humans can naturally identify cause-and-effect relationships in everyday situations, correlations between variables can be misleading. Studies that measure the ability to identify causal relationships directly from factual information are rather sparsely represented.

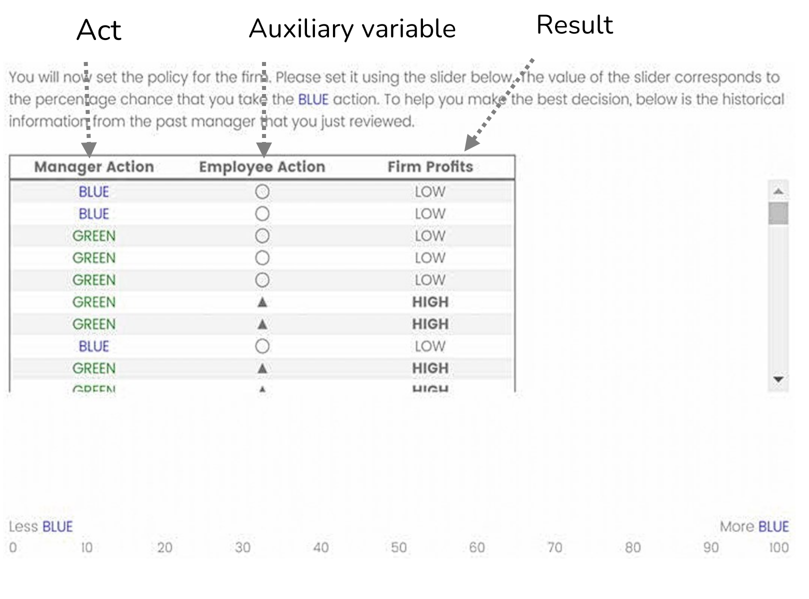
Taxes rates → Taxes base→ budget size

Taxes rates → Taxes base→ budget size

Motivation to earn more

Illustration 1: two different DGP

Illustration 2: Kendall… (2022), Screenshot



Kendall C. W., Charles C. Causal narratives., 2022. paper details

Metrics

- Deviation from optimal policy;
- Degree of confidence

Issues

- Ecological validity (nudging default option)
- Unintuitive task (huge dataset)
- Only one linear DGP pattern.

Research Questions:

- Can the persuasive power of false narratives be counteracted with factual information?
- Do different auxiliary variables, with their corresponding narratives, induce opposite deviations from the rational policy for the same underlying DGP?

literature

As a result, one of the pilot works (Kendall, Charles, 2022) was specifically dedicated to this theme and conducted only in 2022. In this work, authors investigate whether the persuasive power of false narratives could impact on subject's inference. We establish our own framework allowing us to assess this ability directly. The framework involves a three-variable taxonomy on graphs of all possible data generation processes (Eberhardt, 2017) and an approach for incorporating narratives into the task (Kendall, Charles, 2022). Our approach extends the former by using not only linear DGP pattern, and thus tracking dependency between accuracy and complexity (Oprea, 2020).

Task design

Within the task, we offer the subjects to solve the inverse problem repeatedly: observe data and identify the underline DGP. In each round, subjects are given sets of data, divided into three columns with 16 rows each. Each value in each column is the result of a random choice of the system, but the frequency of occurrence depends only on the mechanism assigned to each column. Each pair of columns may be dependent or independent through their mechanisms, which is unknown to the participants. We provide participants with two ways of interacting with data. First, subjects can generate an alternative data table where we are simulating experimental intervention and setting all values of one of the variables to 1. Another way to interact with the task is by filtering each variable by its value.

Link to demo task with instructions:
http://behavior.akadempark.com:8005/demo/data_to_dgp

Illustration 3: task Interfaces

Task design: DGP

This approach allows for the complete identification of the DGP from the data (Hyttinen, Eberhardt, Hoyer, 2013) while avoiding nudging participants toward any solution.

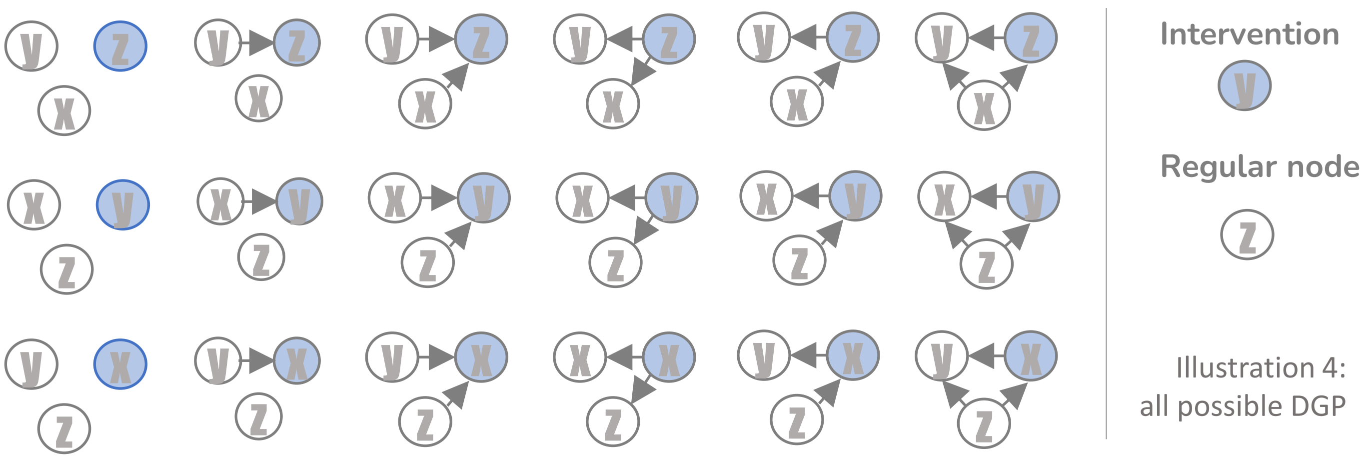


Illustration 4: all possible DGP

Main research question:

In which extent subjects are able to identify causal relationships (involves a three-variable taxonomy on graphs) directly from factual information given various DGP?

- heterogeneity?
- Observational or interventional nature?
- Influence of the complexity of various DGP?

Other questions:

Does feedback impact?
Does narratives impact?
What kind of information (intervention or observational) exploited by subjects more often?

Two stage Design of experiment summary:

For the purpose of analyzing what has a crucial impact on the performance, we also track the influence of narratives. To do this, the framework uses a two-stage design. The first group of participants reports to us their versions of DGP structures, where all answers have corresponding originals. The second stage is identical, except the next group of participants is provided with information about previously reported opinions.

Design of experiment:

control group (13 subjects)*:

- Repeated real effort task with binary variables (6 classes of DGP & 25 unique DGP tasks)
- One round one DGP one task with 3 var.
- 18 round in total (first 6 without feedback and 12 with)
- 2 ways to interact with a data (generate interventional sample on one variable & filter each variable by value)
- Incentivized accuracy & incentivized beliefs
- Collecting reported solutions and confidence

Metrics:

- Measure of accuracy
- Degree of confidence

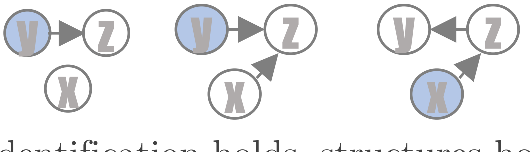
Pilot:

- Same as control (8 subj.*)
- 8 instead of 18 rounds

*already collected

treatment group: (0 subj)

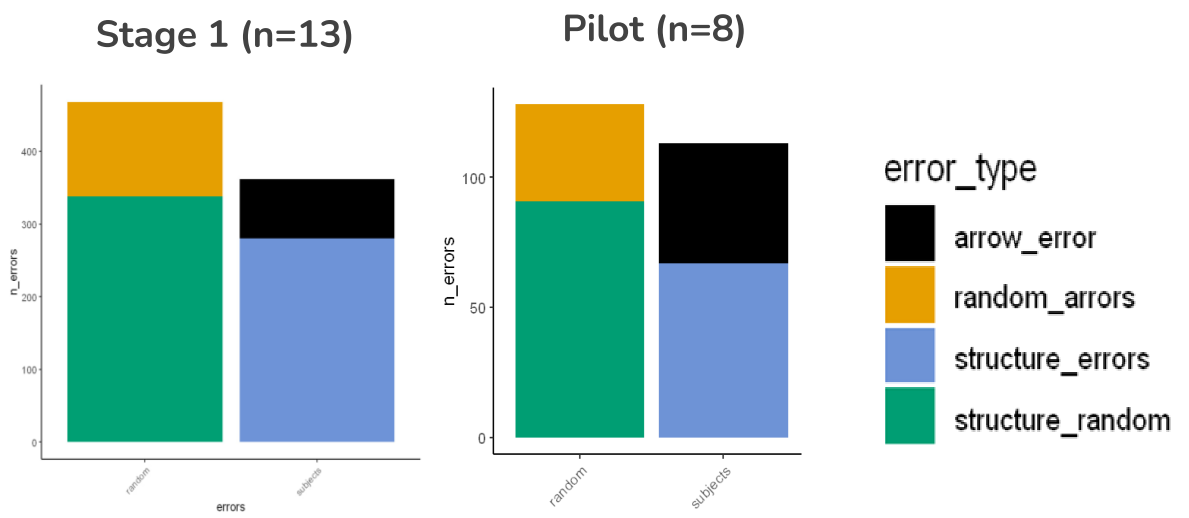
- Same as control
- The node of intervention different:



- Identification holds, structures hold

Brief result summary: Accuracy better than random but far from Ideal

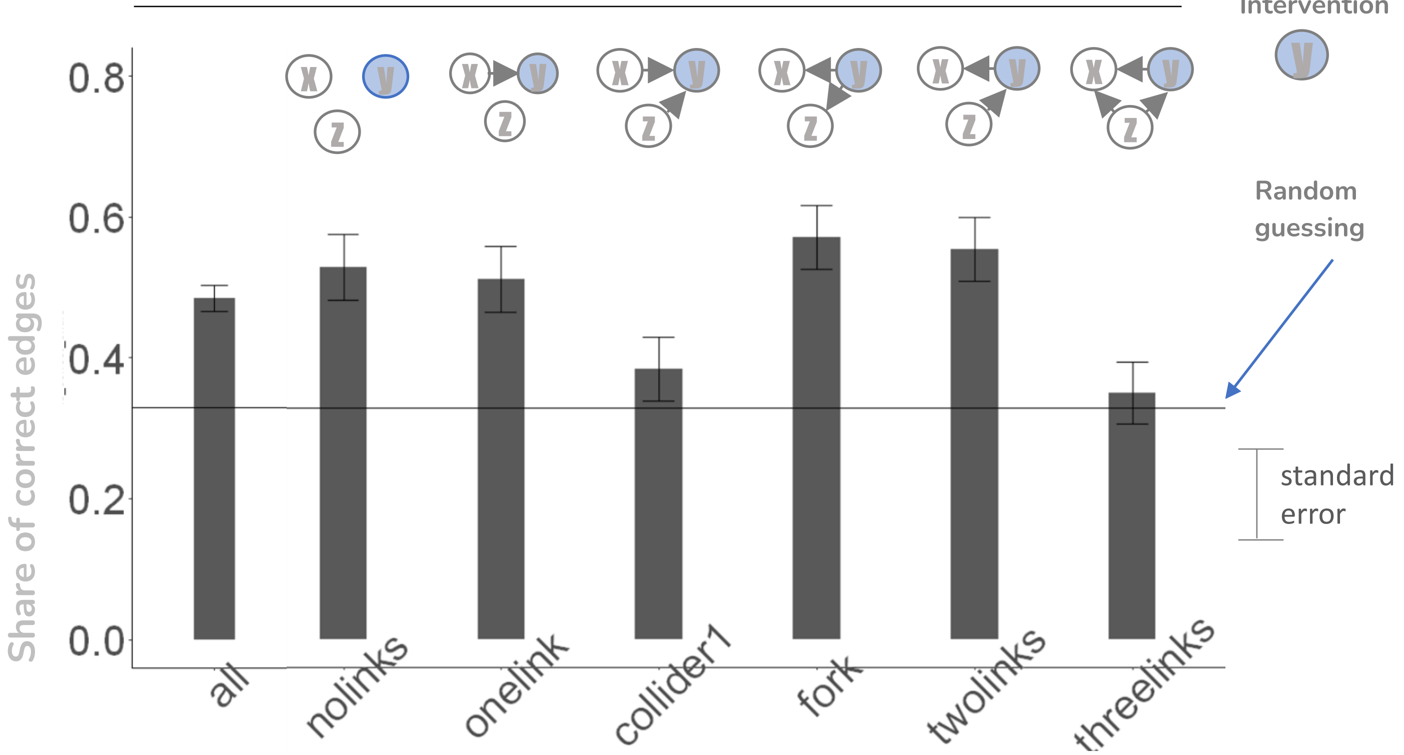
Preliminary, we can say that the common expression correlation does not mean causation is realized to a high degree: people guess the original DGP 10-20 percent better than randomly



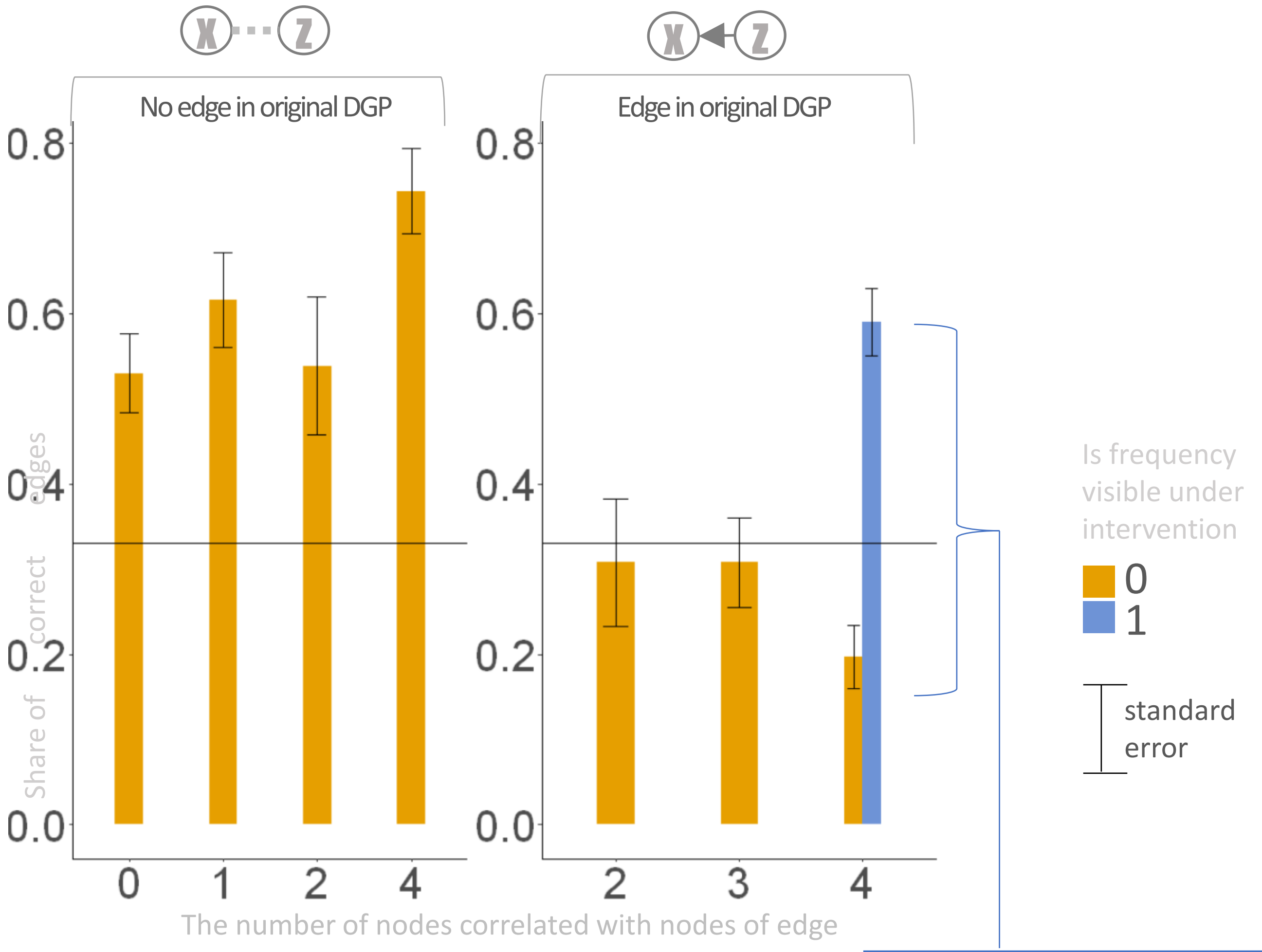
Results: correct answers by DGP shows no pattern (n=13)

After pilot runs, we can see that different DGP lead different accuracy but not according to increasing standard complexity

# edges	0	1	2	2	2	3
# correl.	0	1	2	3	3	3



Instead of standard complexity rules - contrast pattern (n=13)



No edge in original DGP

It seems that “contrast” (graph on the left) helps to participants pefform better: the more the remaining nodes in the graph are connected the better the lack of links is seen on targeted empty place (no edge)

Edge in original DGP

And on case of no contrast (right graph) the opposite is true, the more the remaining nodes in the graph are connected the worse the connection is seen

Intervention contrast

Finally, edges whose outgoing node is an intervention node are also much easier to distinguish, even where there are many links in the graph (again due to the contrast)

References:

Kendall C. W., Charles C. Causal narratives. National Bureau of Economic Research, 2022.

Eberhardt F. Introduction to the foundations of causal discovery. International Journal of Data Science and Analytics. 2017

Oprea R. What makes a rule complex? American economic review. – 2020

Hyttinen A, Eberhardt F, Hoyer PO. Experiment selection for causal discovery. Journal of Machine Learning Research. 2013