## Annotation instructions for cause and effect annotation

In this task, your goal is to find and label cause and effect relationships between variables in documents. If a document describes a cause and effect relationship between X and Y, you should be able to describe the relationship between X and Y using the sentence "X effects Y."

Examples:

The following sentence says that the variable **rice yield** affects the variable **out-migration**

* *The crop-wise analysis shows that a 1 per cent decline in* ***rice yield*** *leads to a nearly 2 per cent increase in the rate of* ***out-migration*** *from a state.*

The following sentence says that the variable **rainfall** affects the variable **remittances**

* *The effect of* ***rainfall*** *on* ***remittances*** *is significantly positive*

The following sentence says that the variable **net inward migration** affects the variable **female labor market**

* ***Net inward migration*** *has strong negative effects on the* ***female labor market***

The following sentence says that **bad weather** affects **risk tolerance** and also says that **bad weather** affects **individual mood**

* Bad weather depresses **individual mood** and **risk tolerance** (i.e. voters are more likely to vote for the candidate who is perceived to be less risky)

If it helps, you can think of a variable as something which would go into a regression.

### How much text to label as a variable?

Sometimes you will encounter cases where the boundaries of a variable can be specified in different ways. For example, the following sentence asserts that inflation affects purchasing power, inflation affects middle-class household purchasing power and inflation reduces Indian middle-class household purchasing power. In these cases, please label the longest-possible version of the variable. For instance, you should label "Indian middle-class household purchasing power.

* *Inflation reduces Indian middle-class household purchasing power*

### Confusing cases

You will encounter some ambiguous cases when reviewing abstracts. Here is how to handle them.

### When to label causal links:

* Sometimes a single phrase will denote several causal links. In this case label each causal link in the sentence. Example:
  + “*rain causes better mood and higher voter turnout*”
    - This sentence contains a causal link between "**rain**" and "**mood**" and also between "**rain"** and "**higher voter turnout**"
* Sometimes variables will be mentioned more than once. Label every single instance where the text mentions a variable pair with a causal link. Example:
  + "*We found that aid causes corruption. Because we found that aid causes corruption we suggest policymakers carefully consider foreign aid.*"
    - In this case, label causal links in each sentence which describes a link between aid and corruption

### When not to label causal links:

* If the document describes a causal relationship from somewhere else in the economics literature, do not label it. You should only record relationships that are stated explicitly in the paper.
  + Example: "Much work in this area has shown that rain affects mood"
* If the text describes a relationship that is not causal do **not** record a causal link. This is called a null relationship.
  + Example: If the abstract says "*We find that salt intake does not cause obesity.*" In this case, do not label anything because there is no explicit causal relationship.
* If you can't tell which variable causes a change in which variable, then don't label it.
  + This will happen if a paper asserts that variables are associated or correlated but does not make claims about cause and effect.
  + Examples**:** “*rain is associated with mood*” or “*rain is correlated with mood*”
    - Notes:
      * We did not include this b/c it is more complex and the simple thing seems to work OK
* Don't label causal links that are not stated explicitly in the text itself. These are implicit and should not be labeled. Examples of implicit assertions:
  + “*This paper studies the effect of rain on mood*.”
    - Do not label this because the paper does not explicitly say that rain affects mood.
  + "*We use exogenous variation in rain to examine the effect of mood on voting*."
    - This assertion is implicit because you can use your knowledge of economics to infer that mood may affect voting, but the text does not explicitly say that mood affects voting.
  + *"The effect was stronger for men."* 
    - Don't label this because the variable (effect of what?) is not stated in the text itself. If you have to look in some other portion of the document to understand what the variable refers to then the causal link is implicit and should not be labeled.
* Sometimes variables are mentioned without a clear connection to other variables. Do not label variables in these cases because there is no stated relationship. Example:
  + "*We control for time.*"
    - In this case, do not label anything because there is no explicit causal relationship. Time implicitly affects some variables, but the connection is not stated explicitly.