

Explanation

Establishing Cause and Effect: Three Conditions

1. Temporal Relationship

Q: Does X precede Y?

A: Empirical evidence in time series data
Logic in cross-sectional data

2. Constant Conjunction

Q: Is X related to Y?

A: Bivariate Crosstabulation

3. Absence of Alternative Explanations

Q: Is X spuriously related to Y?

A: Three-way crosstabulation

Testing for Spuriousness

The Role of Control Variables

Q: Is X truly a cause of Y or are X and Y related because of a third variable Z?

If X is truly a cause of Y then we should be able to provide evidence that X and Y are related regardless of all plausible third variables Z_i .

If X is not a cause of Y (or in other words, if X is spuriously related to Y), then we should be able to find a plausible third variable Z which causes the relationship between X and Y to disappear when we take Z into account.

Five Common Control Effects (The 5 “R’s”)

1. Replicating Effect

No effect is observed. The original relationship is replicated in each of the subtables.

2. Reducing Effect

The original relationship is reduced or eliminated by the introduction of the control variable.

3. Revealing Effect

The original relationship “appears” when the control variable is introduced.

4. Refining Effect

The original relationship is shown to hold up under some conditions (values of the control variable) but not others

5. Reversing Effect

The original relationship is reversed at least in under some conditions when a control is introduced.

EXAMPLES

Replicating

Original relationship:

60	30
60%	30%
40	70
40%	70%

Relationship after Controlling for Z:

Z = 1		Z=2	
30	15	30	15
60%	30%	60%	30%
20	35	20	35
40%	70%	40%	70%

EXAMPLES

Reducing

Original relationship:

60	30
60%	30%
40	70
40%	70%

Relationship after Controlling for Z:

Z = 1		Z=2	
30	15	30	15
60%	30%	60%	30%
20	35	20	35
40%	70%	40%	70%

EXAMPLES

Revealing

Original relationship:

50	50
25%	25%

50	50
25%	25%

Relationship after Controlling for Z:

Z = 1		Z=2	
40	10	10	40
80%	20%	20%	80%
10	40	40	10
20%	80%	80%	20%

EXAMPLES

Refining

Original relationship:

60	30
60%	30%
40	70
40%	70%

Relationship after Controlling for Z:

Z = 1		Z=2	
30	0	30	30
75%	30%	50%	50%
10	40	30	30
25%	100%	50%	50%

EXAMPLES

Reversing

Original relationship:

60	30
60%	30%
40	70
40%	70%

Relationship after Controlling for Z:

Z = 1		Z=2	
60	0	0	30
100%	0%	0%	100%
0	70	40	0
0%	100%	100%	0%

Original relationship between X and Y:

A	B
C	D

Relationship after Controlling for Z:

Z = 1		Z=2	
AZ ₁	BZ ₁	AZ ₂	BZ ₂
CZ ₁	DZ ₁	CZ ₂	DZ ₂

With these 8 cell entries we can reconstruct any combination of the three variables:

The bivariate relationship between X and Z:

The bivariate relationship between Y and Z

The 3-way crosstabulation of Y and Z controlling for X.

The 3-way crosstabulation of X and Z controlling for Y.