

# Simpson's paradox

**Simpson’s paradox**, also called **Yule-Simpson effect**, in statistics, an effect that occurs when the marginal association between two categorical variables is qualitatively different from the partial association between the same two variables after controlling for one or more other variables. Simpson’s paradox is important for three critical reasons. First, people often expect statistical relationships to be immutable. They often are not. The relationship between two variables might increase, decrease, or even change direction depending on the set of variables being controlled. Second, Simpson’s paradox is not simply an obscure phenomenon of interest only to a small group of statisticians. Simpson’s paradox is actually one of a large class of association paradoxes. Third, Simpson’s paradox reminds researchers that causal inferences, particularly in nonexperimental studies, can be hazardous. Uncontrolled and even unobserved variables that would eliminate or reverse the association observed between two variables might exist.

## Illustration

Understanding Simpson’s paradox is easiest in the context of a simple example. Suppose that a university is concerned about sex bias during the admission process to graduate school. To study this, applicants to the university’s graduate programs are classified based on sex and admissions outcome. These data would seem to be consistent with the existence of a sex bias because men (40 percent were admitted) were more likely to be admitted to graduate school than women (25 percent were admitted).

To identify the source of the difference in admission rates for men and women, the university subdivides applicants based on whether they applied to a department in the natural sciences or to one in the social sciences and then conducts the analysis again. Surprisingly, the university finds that the direction of the relationship between sex and outcome has reversed. In natural science departments, women (80 percent were admitted) were more likely to be admitted to graduate school than men (46 percent were admitted); similarly, in social science departments, women (20 percent were admitted) were more likely to be admitted to graduate school than men (4 percent were admitted).

Although the reversal in association that is observed in Simpson’s paradox might seem bewildering, it is actually straightforward. In this example, it occurred because both sex and admissions were related to a third variable, namely, the department. First, women were more likely to apply to social science departments, whereas men were more likely to apply to natural science departments. Second, the acceptance rate in social science departments was much less than that in natural science departments. Because women were more likely than men to apply to programs with low acceptance rates, when department was ignored (i.e., when the data were aggregated over the entire university), it seemed that women were less likely than men to be admitted to graduate school, whereas the reverse was actually true. Although hypothetical examples such as this one are simple to construct, numerous real-life examples can be found easily in the social science and statistics literatures.

## Definition

Consider three random variables *X*, *Y*, and *Z*. Define a  $2 \times 2 \times K$  cross-classification table by assuming that *X* and *Y* can be coded either 0 or 1, and *Z* can be assigned values from 1 to *K*.

The marginal association between *X* and *Y* is assessed by collapsing across or aggregating over the levels of *Z*. The *partial association* between *X* and *Y* controlling for *Z* is the association between *X* and *Y* at each level of *Z* or after adjusting for the levels of *Z*. Simpson’s paradox is said to have occurred when the pattern of marginal association and the pattern of partial association differ.

Various indices exist for assessing the association between two variables. For categorical variables, the odds ratio and the relative risk ratio are the two most common measures of association. Simpson’s paradox is the name applied to differences in the association between two categorical variables, regardless of how that association is measured.

## Association Paradoxes

Association paradoxes, of which Simpson’s paradox is a special case, can occur between continuous (a variable that can take any value) or categorical variables (a variable that can take only certain values). For example, the best-known measure of association between two continuous variables is the correlation coefficient. It is well known that the marginal correlation between two variables can have one sign, whereas the partial correlation between the same two variables after controlling for one or more additional variables has the opposite sign.

Reversal paradoxes, in which the marginal and partial associations between two variables have different signs, such as Simpson’s paradox, are the most dramatic of the association paradoxes. A weaker form of association paradox occurs when the marginal and partial associations have the same sign, but the magnitude of the marginal association falls outside of the range of values of the partial associations computed at individual levels of the variable(s) being controlled. These have been termed amalgamation or aggregation paradoxes.

## Problem of Causality

When confronted with a reversal paradox, it is natural to ask whether the marginal or the partial association is the correct description of the relationship between two variables. Assuming that the relationships among the variables in one’s sample mirror those of the population from which the sample was drawn, then the usual statistical answer is that both the marginal and partial associations are correct. Mathematically, there is nothing surprising about a reversal in the direction of the marginal and partial associations. Furthermore, in an analysis, such as the one presented previously, the reversal of the marginal and partial associations is easily understood once the role of the control variable is understood.

If social scientists were merely interested in cataloging the relationships that exist among the variables that they study, then the answer given previously might be sufficient. It is not. Often, social scientists are interested in understanding causal relationships. In the example given previously, one might be interested in knowing whether the admissions process is biased toward males, as the marginal association might suggest, or biased toward females, as the partial association might suggest. This is the real dilemma posed by Simpson’s paradox for the researcher. It is problematic in two ways.

First, the statistical analysis provides no guidance as to whether the marginal association or the partial association is the spurious relationship. Based on knowledge of graduate admissions, it is reasonable to conclude that the marginal relationship in this example is spurious because admissions decisions are made by departments, not by universities. Substantive information guides this judgment, not the statistical analysis. It might be tempting to conclude, as some authors do, that the marginal association is always spurious. Certainly, that is the impression that is given by much of the published work on Simpson’s paradox. Indeed, some authors characterize Simpson’s paradox as a failure to include a relevant covariate in the design of a study or in the relevant statistical analysis. Unfortunately, this simple answer is inadequate, because it is possible to construct examples in which the partial association is the spurious one. Second, the field of statistics provides limited assistance in determining when Simpson’s paradox will occur. Particularly in nonrandomized studies, there might exist uncontrolled and, even more dangerously, unobserved variables that would eliminate or reverse the association observed between two variables. It can be unsettling to imagine that what is believed to be a causal relationship between two variables is found not to exist or, even worse, is found to be opposite in direction once one discovers the proper variable to control.

## Avoiding Simpson’s Paradox

Although it might be easy to explain why Simpson’s paradox occurs when presented with an example, determining when Simpson’s paradox will occur is more challenging. In experimental research, in which individuals are randomly assigned to treatment conditions, Simpson’s paradox should not occur, no matter what additional variables are included in the analysis. This assumes, of course, that the randomization is effective and that assignment to treatment condition is independent of possible covariates. If so, regardless of whether these covariates are related to the outcome, Simpson’s paradox cannot occur. In nonexperimental, or nonrandomized, research, such as a cross-sectional study in which a sample is selected and then the members of the sample are simultaneously classified with respect to all of the study variables, Simpson’s paradox can be avoided if certain conditions are satisfied. The problem with nonexperimental research is that these conditions will rarely be known to be satisfied a priori.

## History

Given the nature of the phenomenon, perhaps it is only fitting to discover that British statistician Edward Simpson neither discovered nor claimed to have discovered the phenomenon that now bears his name. In his classic 1951 paper, Simpson pointed out that association paradoxes were well known prior to the publication of his paper. Indeed, the existence of association paradoxes with categorical variables was reported by British statistician George Udny Yule as early as 1903. It is for this reason that Simpson’s paradox is sometimes known as the Yule-Simpson effect. It is possible to trace the existence of association paradoxes back even farther in time to British statistician Karl Pearson, who in 1899 demonstrated that marginal and partial associations between continuous variables might differ, giving rise to spurious correlations. Pearson reported that the length and breadth of male skulls from the Paris catacombs correlated .09. The same correlation among female skulls was −.04. After combining the two samples, the correlation was .20. In other words, skull length and breadth were uncorrelated for males and females separately and positively correlated for males and females jointly. Put slightly differently, the marginal association between skull length and breadth was positive, while the partial association between skull length and breadth after controlling for sex was zero.

Not only is Simpson not the discoverer of Simpson’s paradox, but the phenomenon that he described in his 1951 paper is not quite the same as the phenomenon that is now known as Simpson’s paradox. The difference is not critical, but it does reflect the confusion that persists today about what Simpson’s paradox actually is. Some authors reserve the label Simpson’s paradox for a reversal in the *direction* of the marginal and partial association between two categorical variables. Some authors apply Simpson’s paradox to reversals that occur with continuous as well as categorical variables. Still other authors have abandoned the term *Simpson’s paradox* altogether, preferring terms such as *aggregation*, *amalgamation*, or *reversal paradoxes*, which are often defined more broadly than Simpson’s paradox.

Bruce W. Carlson

### TABLE OF CONTENTS

- Introduction
- Illustration
- Definition
- Association Paradoxes
- Problem of Causality
- Avoiding Simpson's Paradox
- History