

2012 1

Statistical Literacy: Confounding

MILO SCHIELD,
Augsburg College
*Director, W. M. Keck Statistical Literacy Project
US Rep., International Statistical Literacy Project
Elected Member: International Statistical Inst.
Webmaster: www.StatLit.org*

May 17, 2012
Slides at www.StatLit.org/pdf/2012Schield8Keene6up.pdf

2012 2

Statistical Literacy

Statistical literacy is **the ability to read and interpret summary statistics in everyday life.**

Statistical Literacy studies

- (1) the relation between statistical associations and causation, and
- (2) the full-range of influences on a statistic or on a statistical association. [Take CARE]

2012 3

Take CARE: Context

The influence of factors **taken into account** by

- data broken out by subgroups in tables and graphs
- averages, ratios and comparisons of averages and ratios
- epidemiological models (cf., deaths attributed to obesity)
- regression models and
- the study design (cf., longitudinal vs. cross-sectional; experiment vs. observational study).

The influence of related factors (confounders) **not taken into account** in the study and **not blocked** by the study design.

2012 4

Controlling for a confounder can DECREASE an association

MN has 3.8 times as much prison expense as ME

| State | Total | # Inmates | Per Inmate |
|-------|--------|-----------|------------|
| MN | \$184M | 4,865 | \$37,825 |
| ME | \$48M | 1,424 | \$33,711 |

MN has 3.4 times as many inmates as ME

MN has 25% more prison expense *per inmate* than ME

2012 5

Controlling for a confounder can NULLIFY an association

MD has 3 times as much prison expense as KS

| State | Total | # Inmates | Per Inmate |
|-------|--------|-----------|------------|
| MD | \$481M | 21,623 | \$22,250 |
| KS | \$159M | 7,148 | \$22,250 |

MD has three times as many inmates as KS

MD has the same prison expense *per inmate* as KS

2012 6

Controlling for a confounder can REVERSE an association

CA has 50% more prison expense than NY

| State | Total | # Inmates | Per Inmate |
|-------|--------|-----------|------------|
| CA | \$2.9B | 136K | \$21,385 |
| NY | \$1.9B | 69K | \$28,426 |

CA has almost twice as many inmates as NY

CA has 25% less prison expense *per inmate* than NY

2012 7

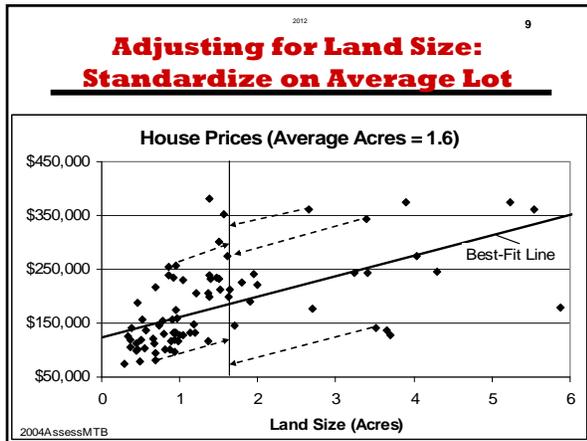
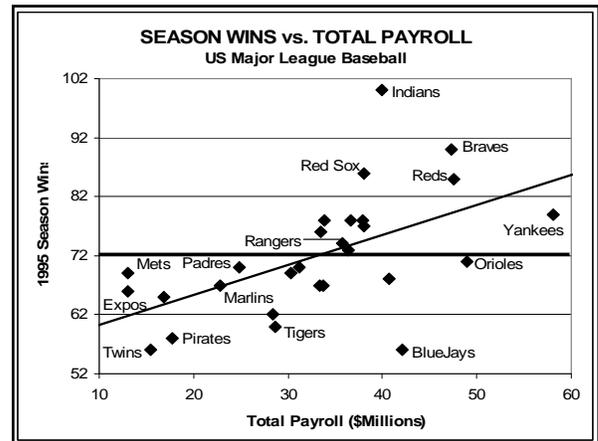
**Controlling for a confounder
can INCREASE an association**

MN has 27% more prison expense than IA

| State | Total | # Inmates | Per Inmate |
|-------|--------|-----------|------------|
| MN | \$184M | 4,865 | \$37,825 |
| IA | \$144M | 5,929 | \$24,286 |

MN has 18% fewer inmates than IA

MN has 56% more prison expense *per inmate* than IA



2012 10

SAT VERBAL SCORES: FLAT

| GROUP | 1981 | 2002 | CHANGE |
|------------------------|-------------------|-------------------|-------------|
| White | 519 (85%) | 527 (65%) | 8 |
| Black | 412 (9%) | 431 (11%) | 19 |
| Asian | 474 (3%) | 501 (10%) | 27 |
| Mexican | 438 (2%) | 446 (4%) | 8 |
| Puerto Rican | 437 (1%) | 455 (3%) | 18 |
| American Indian | 471 (0%) | 479 (1%) | 8 |
| ALL Test takers | 504 (100%) | 504 (100%) | ZERO |

2012 11

**Multivariate Analysis
can be Complex**

To simplify, consider cases with

- a binary outcome,
- a binary predictor and
- a binary confounder.

What are the necessary conditions for nullification or a reversal?

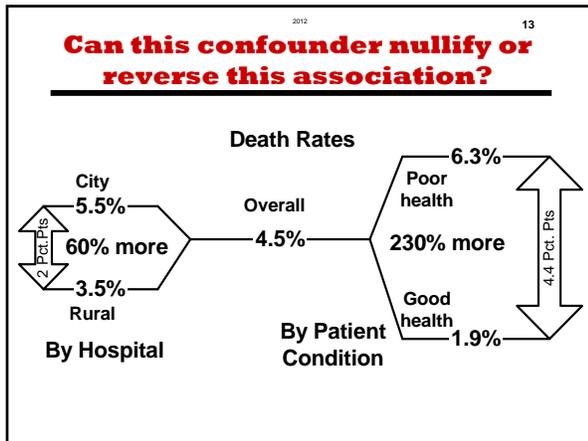
See Schield (1999) and Schield and Burnham (2003)

2012 12

**City Hospital:
Hospital of Death??**

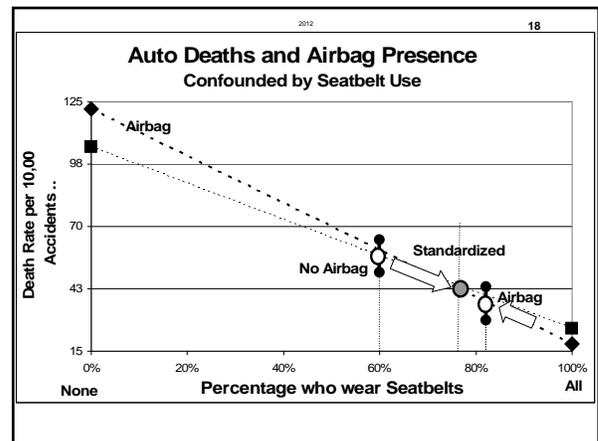
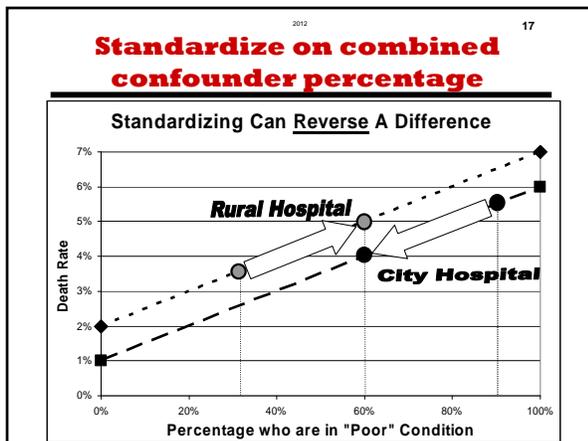
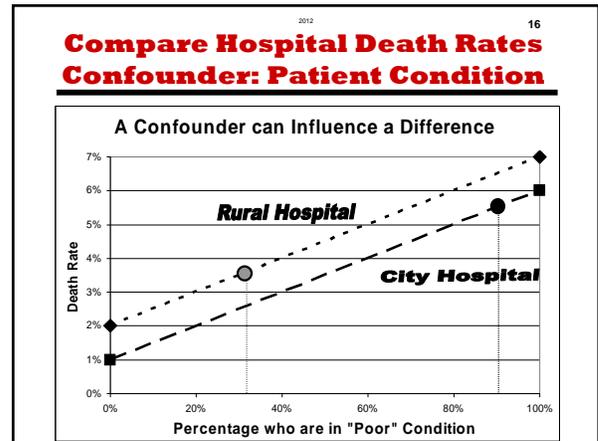
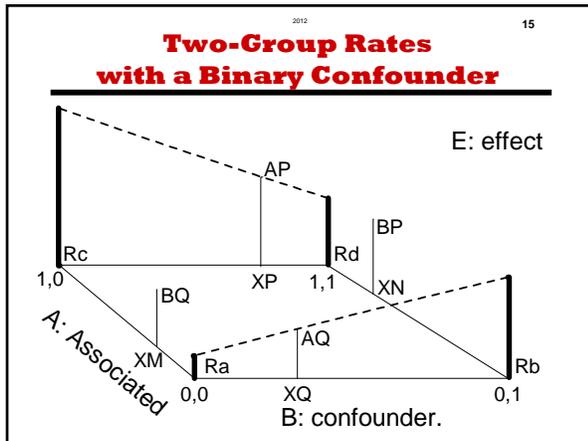
| Hospital | Total | Died | Death Rate |
|----------|-------|------|------------|
| City | 1,000 | 55 | 5.50% |
| Rural | 1,000 | 35 | 3.50% |
| Both | 2,000 | 90 | 4.50% |

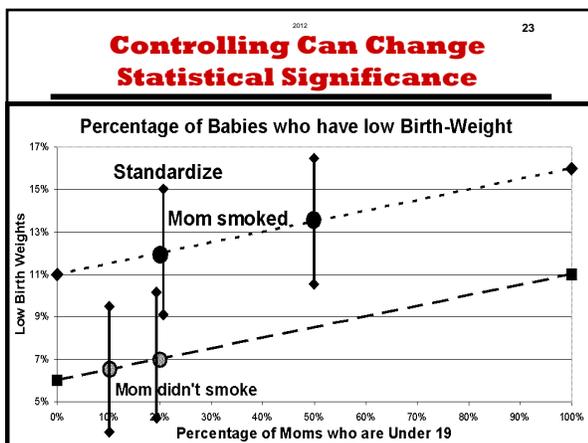
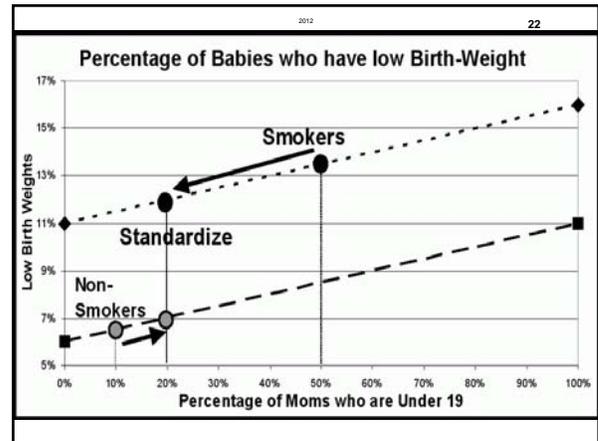
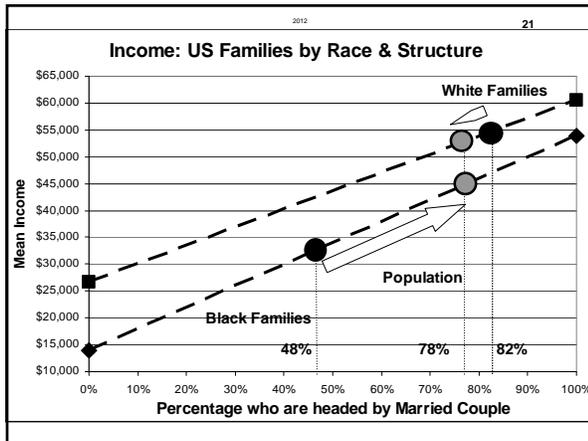
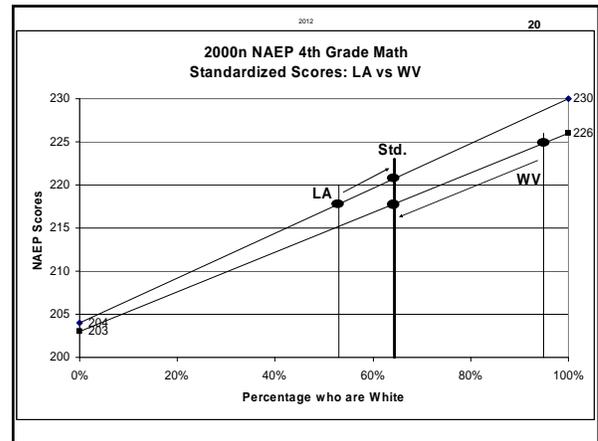
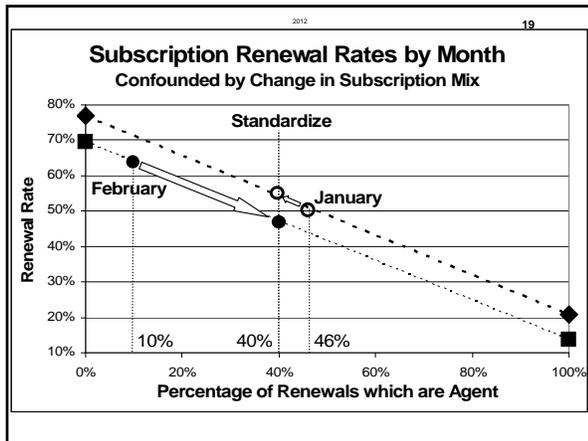
| Condition | Total | Died | Death Rate |
|-----------|-------|------|------------|
| Good | 800 | 15 | 1.90% |
| Poor | 1,200 | 75 | 6.30% |



Confounder Reverses; City Hospital is Better

| Condition | Hospital | Total | Died | Death Rate |
|-----------|----------|-------|------|------------|
| Good | City | 100 | 1 | 1.00% |
| | Rural | 700 | 14 | 2.00% |
| | Total | 800 | 15 | 1.90% |
| Poor | City | 900 | 54 | 6.00% |
| | Rural | 300 | 21 | 7.00% |
| | Total | 1,200 | 75 | 6.30% |





2012 24

Conclusion

Statistical educators must show students how confounders can influence associations and change statistical significance. The failure of educators to do this may be seen as “*statistical negligence.*”

Schild (1999). Simpson's Paradox and Cornfield's Conditions. See www.StatLit.org/pdf/1999SchildASA.pdf.

Schild, Milo (2006). Presenting Confounding and Standardization Graphically. *STATS Magazine*, ASA, Fall 2006, pp. 14-18. Draft at www.StatLit.org/pdf/2006SchildSTATS.pdf.

Schild, Milo (2009). Confound Those Speculative Statistics. *2009 ASA Proceedings of the Section on Statistical Education*. [CD-ROM] 4255-4266. www.StatLit.org/pdf/2009SchildASA.pdf