Algebraic Conditions for Binary Spuriousity

MILO SCHIELD

Augsburg College Department of Business Administration

Director, W. M. Keck Statistical Literacy Project

THOMAS V.V. BURNHAM

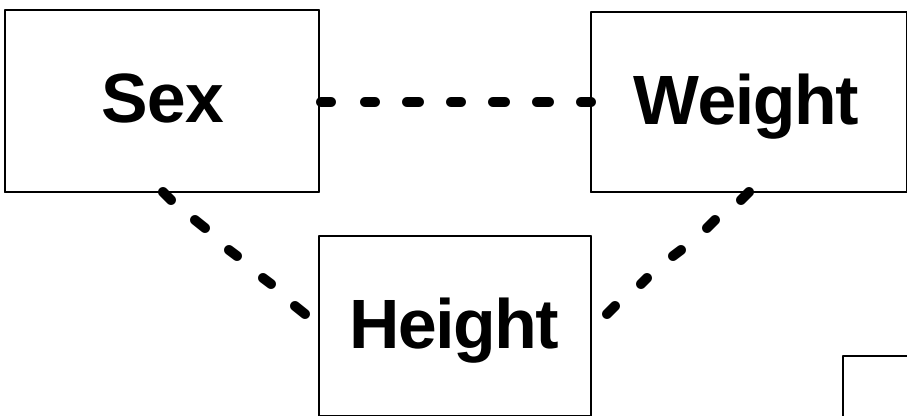
Cognitive Consulting

August 4, 2003

www.augsburg.edu/ppages/~schield **schield@augsburg.edu**

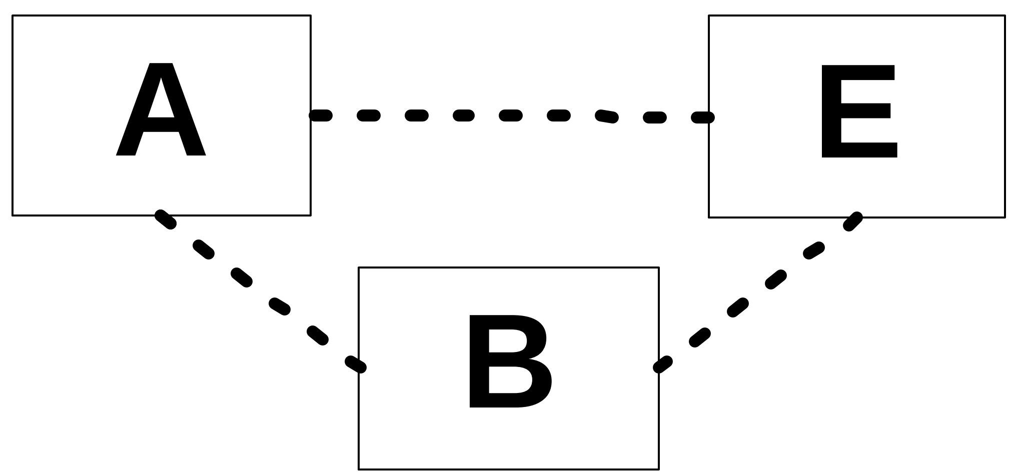
Associations Confounded
No test for Confounding

In observational studies,
associations are often *confounded (tangled up)*.

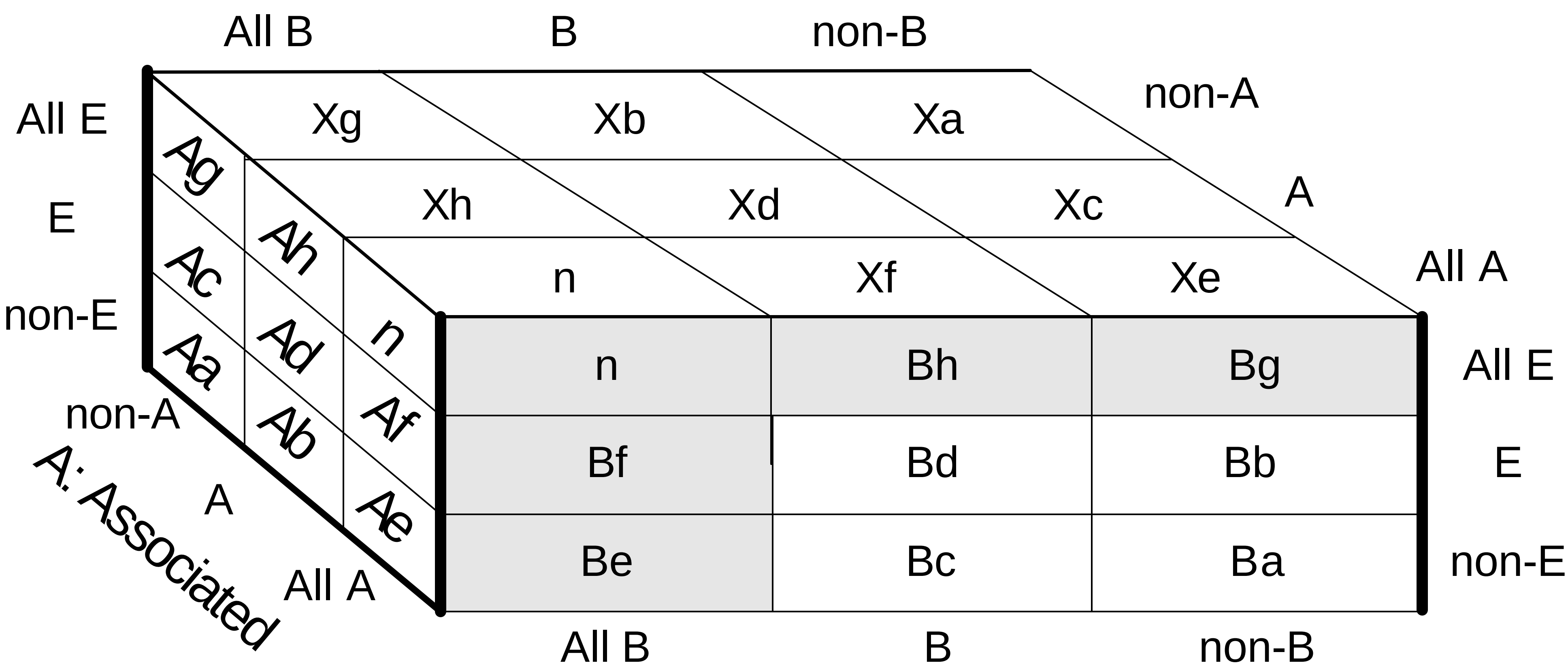


Confounding Factor

E is Effect of interest
A is associated predictor
B is confounder



Categorical Cube: Three Binary Variables

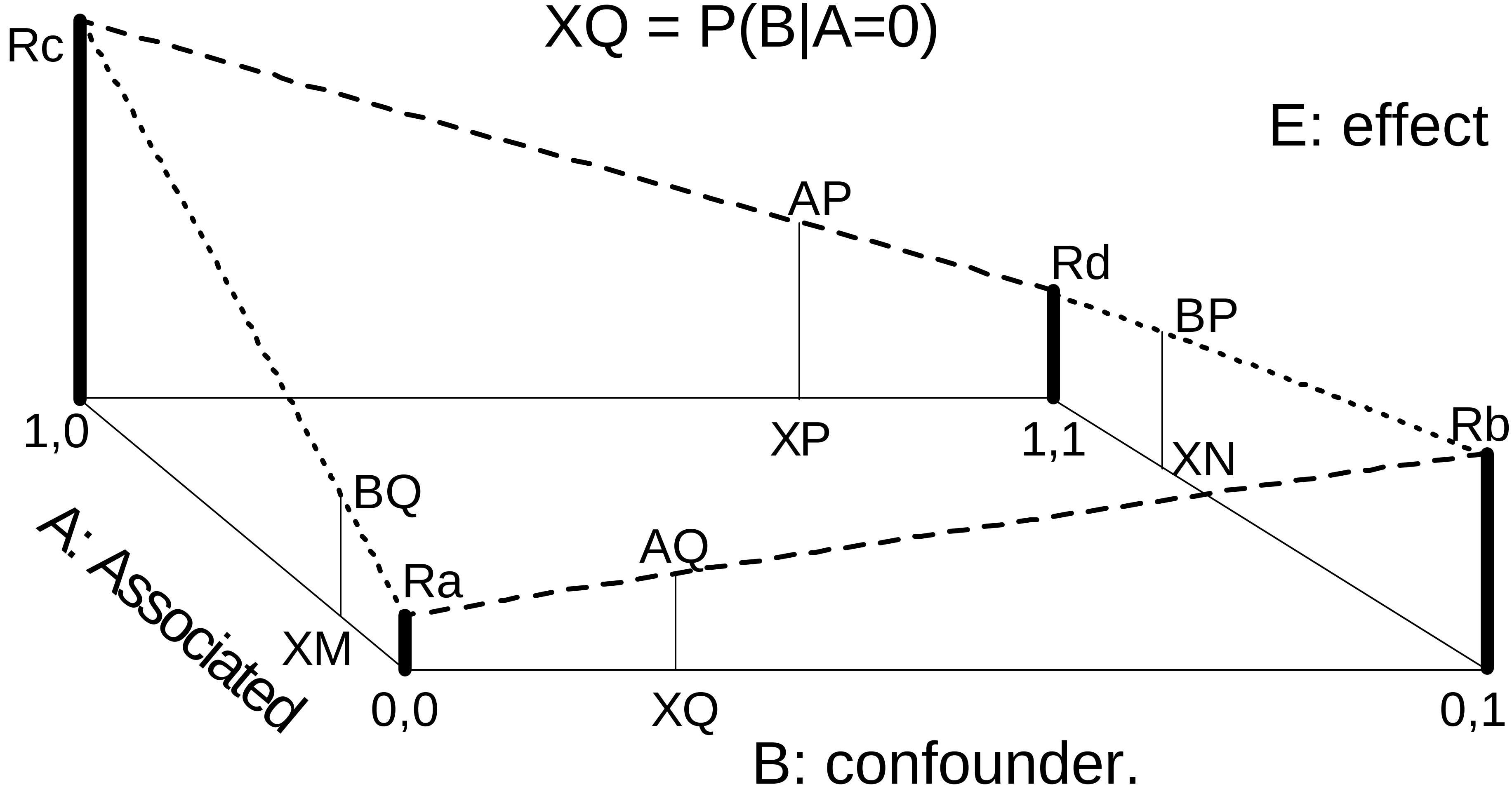


B: confounder.

Quantitative Rate Cube Non-Planar Data

AQ = P(E|A=0,B=XQ) = Rb*XQ + Ra*(1-XQ)

XQ = P(B|A=0)



Criteria for Spuriousity: *A* has "no effect" on *E*

Cornfield & Gastwirth used a cross-A rate equality model:

- ***P(E|A and B) = P(E|B) = P(E|non-A and B)***
- ***P(E|A and non-B) = P(E|non-B) = P(E|non-A and non-B)***

We used two regression models:

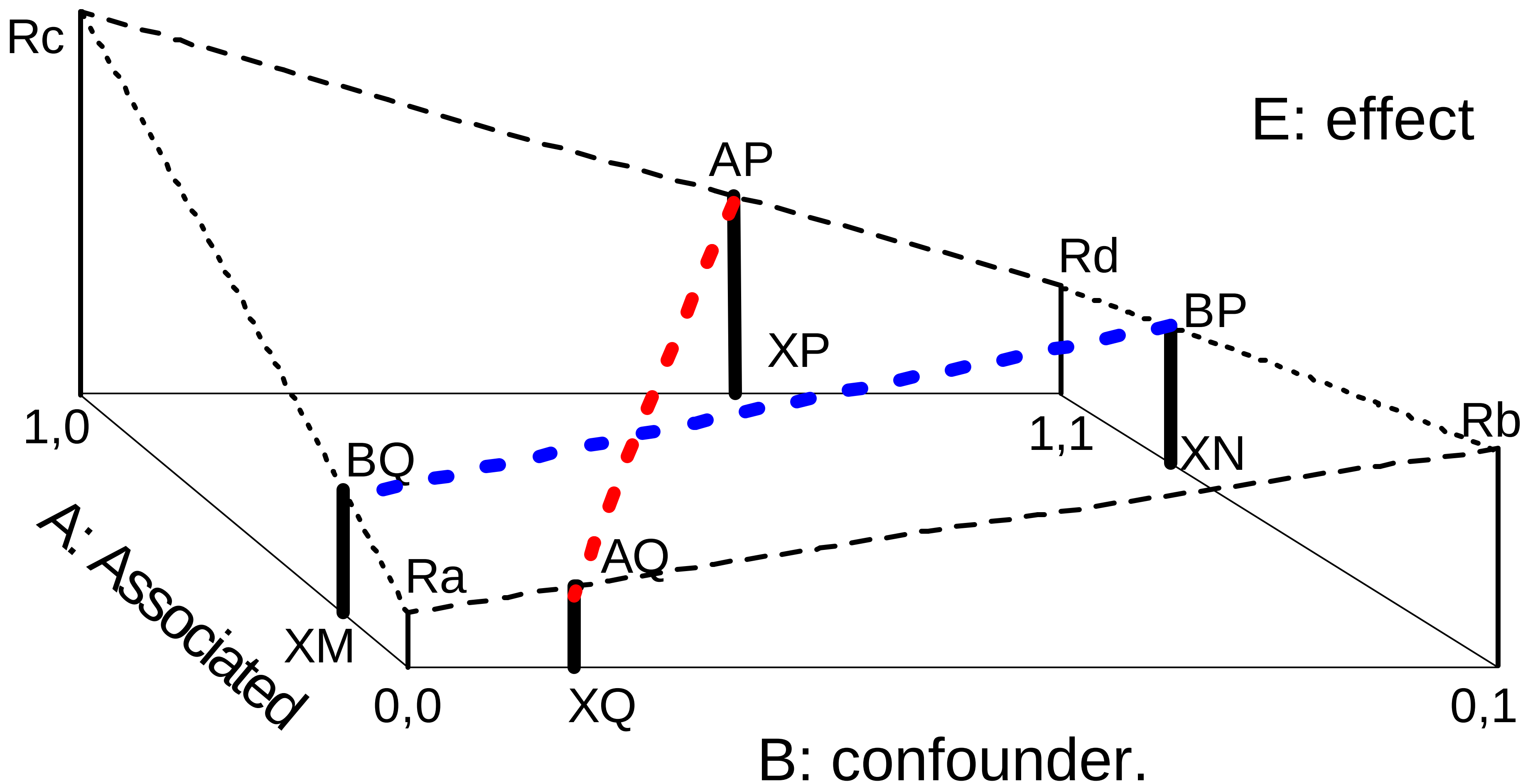
- ***A non-interactive model: E = bo + b1*A + b2*B***
- ***An interactive model: E = b0 + (b1 + b3*B) *A + b2*B***

A-E association is spurious if underlined factor is zero.

As viewed from confounder perspective: B-E

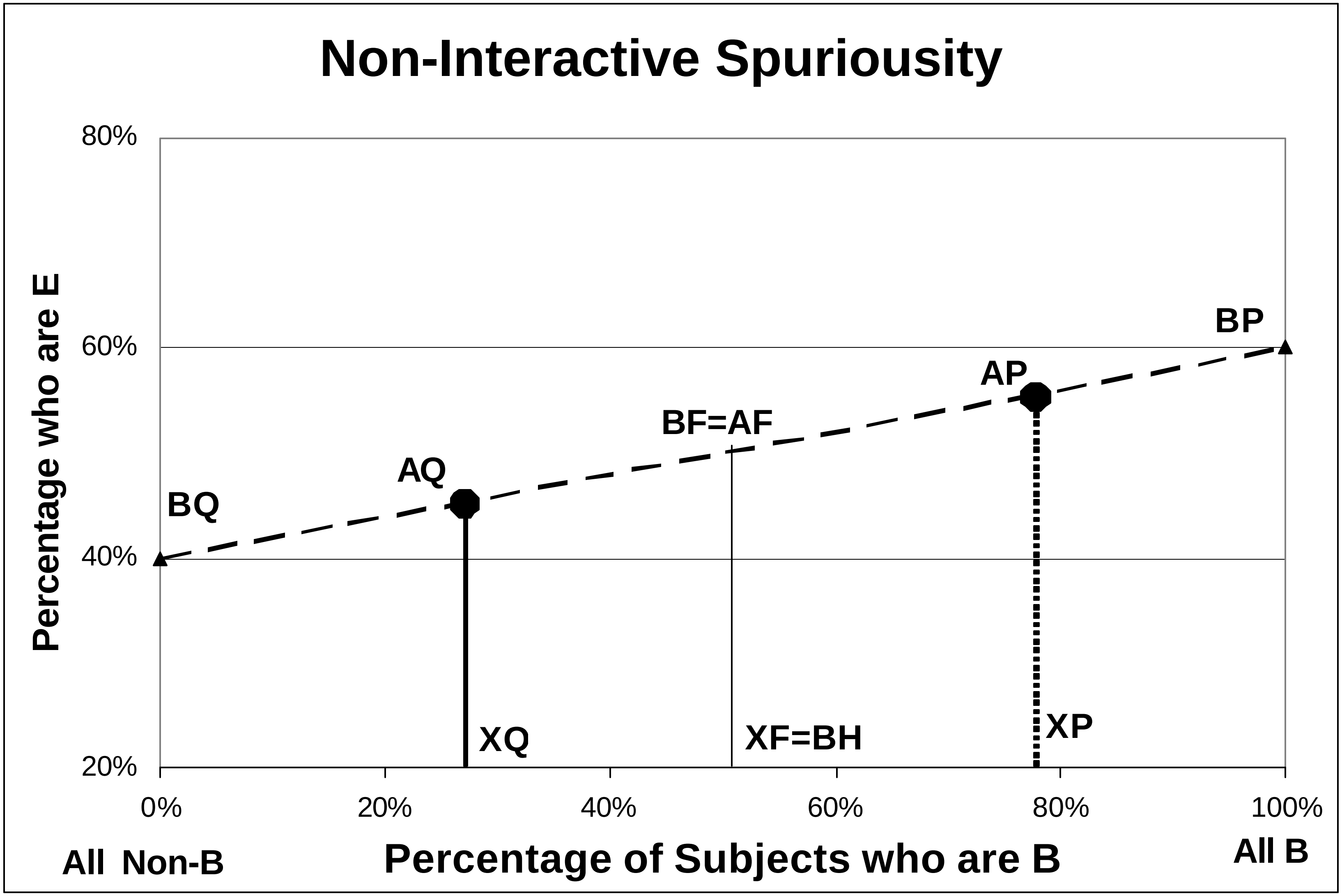
- Non-interactive model: B line || A line
- Interactive model: Rate lines intersect at prevalence of B.

Non-Interactive Model: AP:AQ line and BP:BQ line

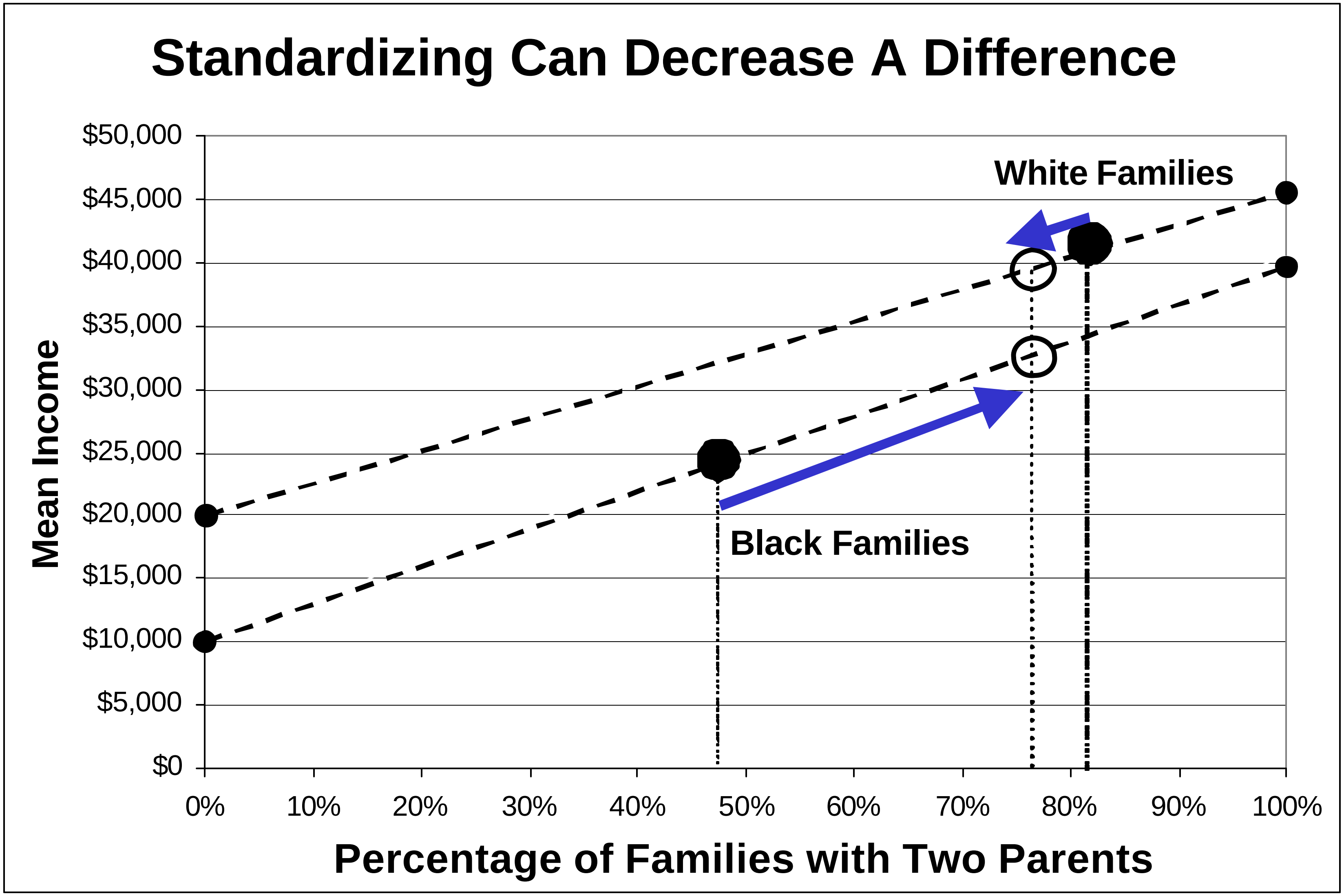


Non-interactive Spuriousity Projected on B:E Face

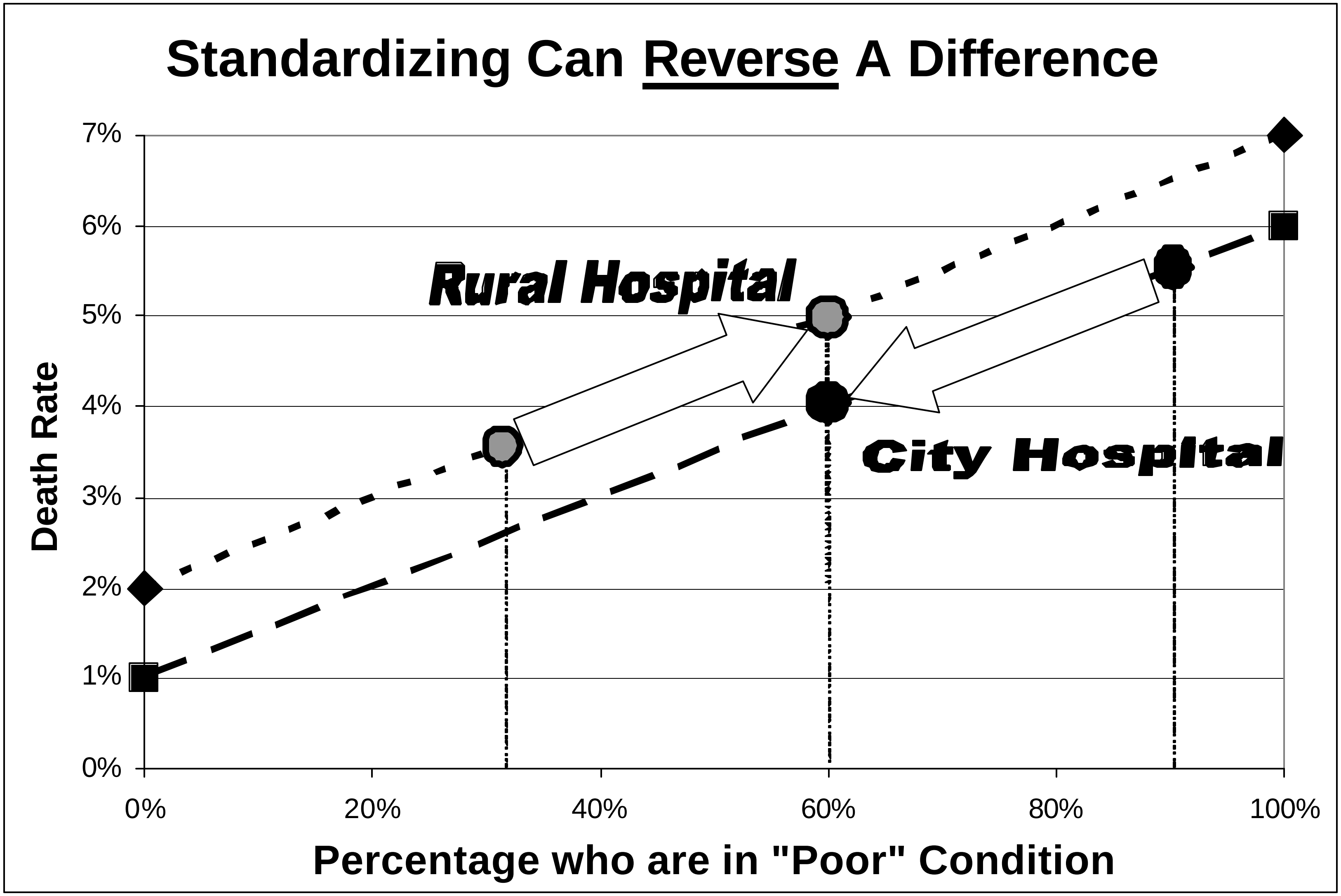
•



Standardizing Shows Influence of Confounder

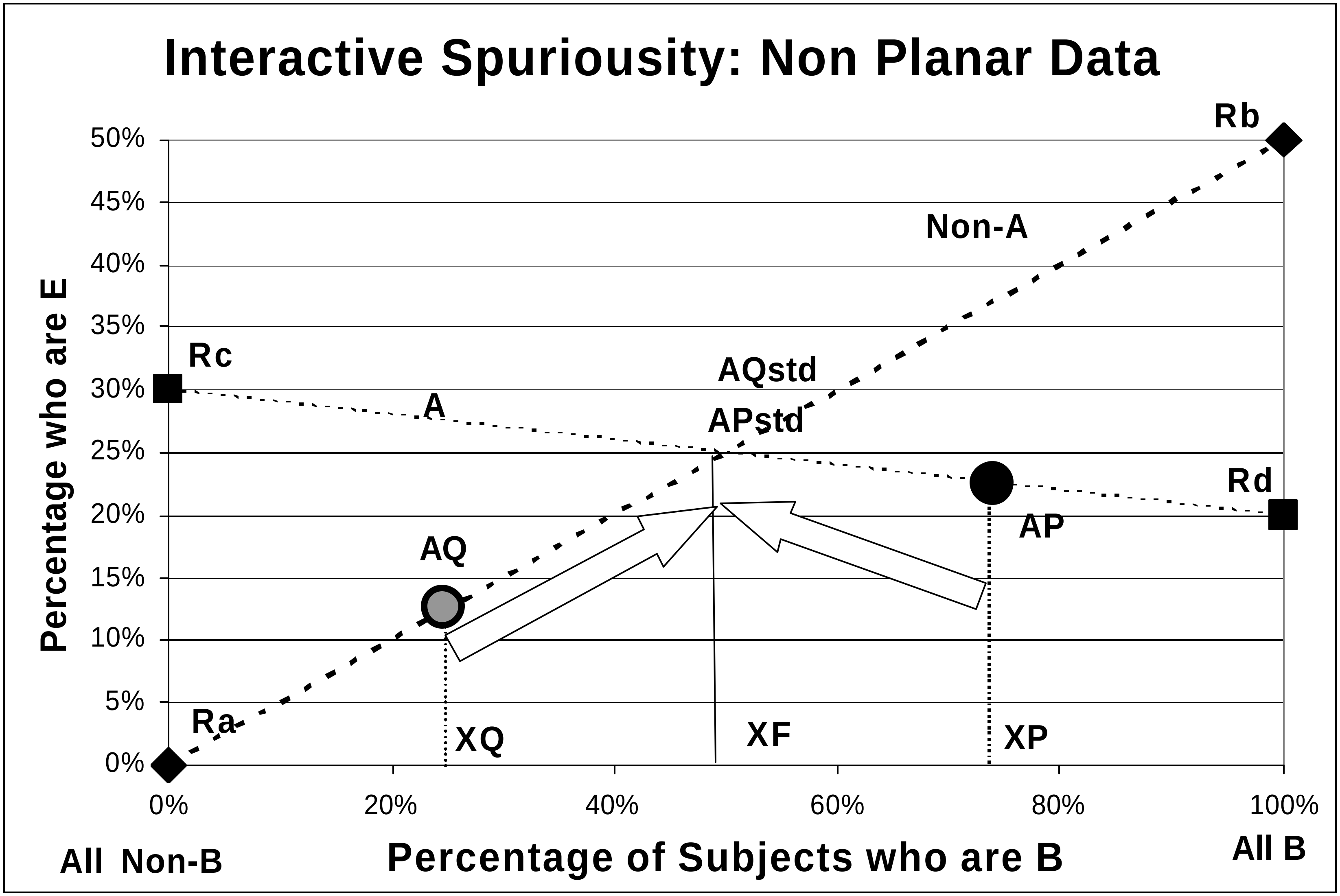


Standardizing Shows Simpson's Paradox



Interactive Spuriousity via Standardizing

•



Spuriousity Results: New Necessary Condition

Gastwirth-Cornfield: RR(E:B) > RR(E:A)

New: RR(E:B) - 1 > [RR(E:A) - 1][P(A)/P(B)]

What cancer-gene effect size is necessary to make association between smoking and cancer spurious?

RR(E:A)=9 for cancer among smokers vs. non.
P(B) = 10%. 10% of adults have a cancer gene
P(A) = 40%. 40% of adults smoke, then

- Gastwirth-Cornfield: RR(E:B) > 9.
- New: RR(E:B) > 33

Conclusions

Spuriousity depends on model.

Cornfield conditions more-generally valid.

Standardizing illustrates interactive model.

Spuriousity conditions for non-interactive and interactive models overlap.

New equations for non-interactive spuriousity.

New inequality for non-interactive model:

$$RR(E:B)-1 > [RR(E:A)-1] \bullet P(A) / P(B)$$