Subject: Re: Odds ratios vs Marginal effects
Posted by Reduced-For(u)m on Sat, 23 Jan 2016 00:03:57 GMT
View Forum Message <> Reply to Message

Gabriel,

 Good question. I think these are two separate issues:

1 - how you report effect sizes: as marginal effects, odds ratios, hazard rates...whatever. This is just a choice of units. It has no real relationship to the correlation/causation question. A regression slope is (often) just another way to summarize the same information in an odd-ratio or a hazard.

Sometimes it might appear that choice of units relates to "causal interpretations", but that is probably because different disciplines use different terminology, have different standards for what constitutes a "causal effect", and tend to use primarily one of the potential effect size measures. So the units used for reporting effect sizes and the causal (or correlational) language used by practitioners who report those effects are often highly correlated, because they are similarly trained.

2 - I think when they say you can't get "causal estimates" from the DHS data, they are simply pointing out that this is observational data, and not the result of some particular experiment with experimentally-assigned treatment groups. From what I understand, this has been standard training in many biomedical fields for a long time. In the social sciences the development of the concept of "Natural Experiment", along with associated methodology, has led to a generation of practitioners trained to believe in both experimental and non-experimental methods for estimating causal relationships (Instrumental Variables, Regression Discontinuity, certain kinds of Difference-in-Difference). Ironically, I think the first Natural Experiment was actually the work of John Snow in Epidemiology*, but these days it is mostly Econ/Poli Sci that think about natural experiments.

That said, in order to use DHS data to do a "Natural Experiment" you usually have to import some sort of external data (an instrument, a policy roll-out, something). So in a sense, I agree that just using pure DHS data usually means estimating "correlations" or (as an idea I'm developing "Deep Correlations**", those purged of obvious and confounding observables) and not causal effects. But it isn't a given that no causal effects can be estimated using DHS data. I would argue that people who say that are really just saying that only experiments can generate causal estimates, and I think that is a rather narrow view of how we conduct statistical inference.

*See "Statistical Models and Shoeleather" by David Freedman for a discussion of Snow's awesome Natural Experiment
**This is not (at least yet) a well-defined or mathematically grounded concept, just an idea I have to distinguish certain kinds of deeply meaningful correlations from other kinds of more superficial correlations.

Thoughts? Reminder: I am not affiliated with the DHS and my responses do not necessarily reflect the views of anyone other than me.