

# How to Accuse the Other Guy of Lying with Statistics

Charles Murray

*Abstract.* We've known how to lie with statistics for 50 years now. What we really need are theory and praxis for accusing someone else of lying with statistics. The author's experience with the response to *The Bell Curve* has led him to suspect that such a formulation already exists, probably imparted during a secret initiation for professors in the social sciences. This article represents his best attempt to reconstruct what must be in it.

*Key words and phrases:* Public policy, regression analysis, lying with statistics.

In 1994, the late Richard J. Herrnstein and I published *The Bell Curve* (Herrnstein and Murray, 1994) and set off an avalanche of editorials, news stories, articles and entire books in response. That avalanche included valuable technical contributions that have moved the debate forward. But much of the reaction that went under the cover of scholarly critique baffled me because it seemed transparently meretricious. These people were too smart and well trained to believe their own arguments, I said to myself, and I spent many hours imagining how they rationalized lying (in my not-disinterested view) about the book's arguments and evidence. But *The Bell Curve* wasn't a unique case. For books on certain high-profile policy issues—Bjorn Lomborg's *The Skeptical Environmentalist* (Lomborg, 1998) is another prominent example—the ordinary rules constraining scholarly debate seem to go out the window. In my more paranoid moments, I envision a secret initiation for newly-appointed assistant professors in the social sciences that goes something like this:

\*

Over the last few decades, a number of books on public policy aimed at a lay readership have advanced conclusions that no socially responsible person can abide, written so cleverly that they have misled many gullible people.

---

Charles Murray is the Bradley Fellow at the American Enterprise Institute for Public Policy Research, 1150 Seventeenth Street, N.W., Washington, D.C. 20036, USA (e-mail: charlesmurray@adelphia.net).

Unfortunately, the people who write such books often call upon data that have some validity, which confronts us with a dilemma. Such books must be discredited, but if we remain strictly within the rules of scholarly discourse, they won't be. What to do? Recall Lawrence Kohlberg's theory of moral development: At the sixth and highest level of morality, it is permissible to violate ordinary ethical conventions to serve a higher good (Kohlberg, 1981). Such is the situation forced upon us by these books. Let me offer six strategies that you may adapt to the specific situation you face.

As you consider these strategies, always keep in mind the cardinal principle when attacking the target book: *Hardly anyone in your audience will have read it.* If you can convince the great majority who never open the book, it doesn't matter that the tiny minority who have read it will know what you are doing.

## #1. THE WHOLE THING IS A MESS

This is a form of softening up, "preparing the battlefield" as the military would put it. The goal is to generate maximum smoke. The specific criticisms need not be central to the target book's argument. They need not even be relevant. All you need to do is to create an impression of many errors, concluding with, "If a sophomore submitted this as a paper in my introductory [insert name of course], I would flunk it."

Samples offer a rich source of smoke. Something is wrong with every sample. Start with that assumption, which has the advantage of being true, seek out that something, and then announce that the data are uninterpretable. If the sample is representative, argue that the

data are outdated. If the sample is recent, argue that it is unrepresentative. If it is both recent and representative, you may be able to get some mileage out of missing data. If the author drops cases with missing data, argue that doing so biases the sample. If instead the author uses imputed values, accuse him of making up data.

Another excellent way to create smoke is to focus on the target book's  $R^2$ 's, which are almost always going to be smaller than 0.5 and often will be around 0.1 or 0.2. The general form of the accusation in this case is, "[The independent variable] that the author claims is so important explains only [ $x$ ] percent of the variance in [the dependent variable]. That means that [ $100-x$ ] percent is the result of other causes. The role of [the author's independent variable] is trivial." Do not let slip that your own published work is based on similarly low  $R^2$ 's.

A third generic way to create smoke is to accuse of the author of choosing the wrong analytical model. The author chose a linear model when it should have been nonlinear. He chose a tobit model instead of a negative binomial model. He used a fixed-effects model instead of a random-effects model. Here the general form of your position is, "Even a first-year graduate student would know better than to use [the target's choice of model] instead of [the preferred model]." Do not be deterred if the results are robust across alternative models. Remember the cardinal rule: Hardly anyone will have read the book, so hardly anyone will know.

## #2. KEEP ADDING INDEPENDENT VARIABLES

Now you are ready to demonstrate that the author is not only incompetent, but wrong. If you have access to data for replicating the target book's analysis, one statistical tool is so uniformly devastating that no critic should be without it: Keep adding independent variables. Don't worry if the new variables are not causally antecedent to the author's independent variables. You can achieve the same result by adding correlated independent variables that are causally posterior. The regression coefficients for the key variables in the target book's analyses will be attenuated and sometimes become statistically insignificant. Technical note: Combine the old and new variables into a single-equation model, not into a multi-equation model. You don't want to give your reader a chance to realize that you're saying that the sun rises because it gets light.

So far, I have given you some tools for fighting statistics with statistics. But remember Frederick Mosteller's dictum that while it is easy to lie with statistics, it is even easier to lie without them. Let me turn

now to refutations of statistical evidence that exploit this profound truth.

## #3. ANY ALTERNATIVE EXPLANATION THAT CAN BE IMAGINED IS TRUE

The first of these ways to fight evidence without evidence calls on the power of the alternative explanatory hypothesis. As every poker player knows, it is not necessary actually to have good cards if you play the hand *as if* you had good cards. Similarly, you can advance competing hypotheses as if they are known to be true, as in this form: "The author fails to acknowledge that [some other cause] can have [the observed effect], invalidating the author's explanation." Technical note: Don't make the beginner's mistake of using "could" instead of "can" in this formulation—a careful reader might notice the implication that the alternative has no evidence to back it up.

## #4. NOTHING IS INNOCENT

If you can persuade your audience that the author of the target book is slanting the data, you cast a cloud of suspicion over everything the author says. Thus the rationale for strategy #4, again happily requiring no evidence: Treat any inconsistency or complication in the target book's interpretation of the data as deliberately duplicitous. Some useful phrases are that the author "tries to obscure..." or "conspicuously fails to mention..." or "pretends not to be aware that..." Here, remember that the more detailed the book's technical presentation, the more ammunition you have: any time the author introduces a caveat or an alternative interpretation in an endnote or appendix, it has been deliberately hidden.

## #5. SOMEONE SOMEWHERE SOMETIME HAS SAID WHAT YOU PREFER TO BE TRUE

Sometimes the target book will use evidence based on a review of the extant technical literature. Such evidence is as easy to attack as the quantitative evidence if you remember "The Rule of One," which is as follows: In a literature in which a large number of studies find  $X$  but even one study finds not- $X$ , and the finding  $X$  is pernicious, you may ignore the many and focus exclusively on the one. Ideally, the target book will not have cited the anomalous study, allowing you to charge that the author deliberately ignored it (see strategy #4). But even if the target book includes the anomalous study in its literature review, you can still treat the one as definitive. Don't mention the many.

A related principle is the “Preferential Option for the Most Favorable Finding,” applied to panel studies and/or disaggregated results for subsamples. If the author of the target book has mentioned the overall results of such a study, find the results for one of the panels or one of the subsamples that are inconsistent with the overall finding, and focus on them. As you gain experience, you will eventually be able to attack the target book using one subsample from an early panel and another subsample from a later panel without anyone noticing.

#### #6. THE JUDICIOUS USE OF THE BIG LIE

Finally, let us turn from strategies based on half-truths and misdirection to a more ambitious approach: to borrow from Goebbels, the Big Lie.

The necessary and sufficient condition for a successful Big Lie is that the target book has at some point discussed a politically sensitive issue involving gender, race, class or the environment, and has treated this issue as a scientifically legitimate subject of investigation (note that the discussion need not be a long one, nor is it required that the target book takes a strong position, nor need the topic be relevant to the book’s main argument). Once this condition is met, you can restate the book’s position on this topic in a way that most people will find repugnant (e.g., women are inferior to men, blacks are inferior to whites, we don’t need to worry about the environment), and then claim that this repugnant position is what the book is about.

What makes the Big Lie so powerful is the multiplier effect you can get from the media. A television news show or a syndicated columnist is unlikely to repeat a technical criticism of the book, but a nicely framed Big Lie can be newsworthy. And remember: It’s not just the public who won’t read the target book. *Hardly anybody in the media will read it either.* If you can get your accusation into one important outlet, you can start a chain reaction. Others will repeat your accusation, soon it will become the conventional wisdom, and no one will remember who started it. Done right, the Big Lie can forever after define the target book in the public mind.

\*

So there you have it: six tough but effective strategies for making people think that the target book is an irredeemable mess, the findings are meaningless, the author is incompetent and devious and the book’s thesis is something it isn’t. Good luck and good hunting.

#### REFERENCES

- HERRNSTEIN, R. J. and MURRAY, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press, New York.
- KOHLBERG, L. (1981). *The Philosophy of Moral Development: Moral Stages and the Idea of Justice*. Harper and Row, San Francisco.
- LOMBORG, B. (1998). *The Skeptical Environmentalist: Measuring the Real State of the World*. Cambridge Univ. Press.