

## Statistics for Cigarette Sellers

Andrew Gelman

### Part 1: The Journalist

Remember *How to Lie with Statistics*? It turns out the author worked for the cigarette companies. Historian Robert Proctor, in his book, *Golden Holocaust: Origins of the Cigarette Catastrophe and the Case for Abolition*, writes:

Darrell Huff, author of the wildly popular (and aptly named) *How to Lie with Statistics*, was paid to testify before Congress in the 1950s and then again in the 1960s, with the assigned task of ridiculing any notion of a cigarette-disease link. On March 22, 1965, Huff testified at hearings on cigarette labeling and advertising, accusing the recent surgeon general's report of myriad failures and "fallacies." Huff peppered his attack with amusing asides and anecdotes, lampooning spurious correlations like that between the size of Dutch families and the number of storks nesting on rooftops—which proves not that storks bring babies but rather that people with large families tend to have larger houses (which therefore attract more storks).

This was all a surprise to me, and I suspect to other statisticians as well. For example, Huff's activities with the cigarette companies are not mentioned on his Wikipedia page (at the time of this writing), nor were they mentioned in a 2005 appreciation of Huff by probabilist J. Michael Steele in the journal *Statistical Science*.

Huff is best known for his classic on statistical communication, but he wrote several other books, including *Pictures by Pete* (1944), *The Dog That Came True* (1946), *How to Take a Chance* (1959), *Score: The Strategy of Taking Tests* (1961), and *How to Lower Your Food Bills* (1963).

It appears he was also working on a book in the late 1960s called *How to Lie with Smoking Statistics*, which the publisher saw "high likelihood of proceeding into print."

In November 1965, a letter was sent to Huff as follows:

Dear Darrell:

Somewhat belatedly, here is the sequel to our telephone conversation of last Friday.

Mr. Nagourney at Macmillan informs me that the changes you made in the contract mailed back to Macmillan are agreeable. Mr. Nagourney is having the contract re-typed, he tells me, for dispatch to you. This should make it feasible for you to proceed with the prescribed changes in your book.

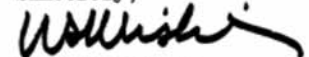
For your information, the relationship between The Tobacco Institute and The Tiderock Corporation has been severed. Accordingly, future discussions of the book should be with Mr. William Kloepfer, Vice President-Public Relations, The Tobacco Institute, 1735 K Street, N.W., Washington, D.C.

All this information comes from the Legacy Tobacco Documents Library (<http://legacy.library.ucsf.edu>), where a search on Huff also turned up a letter from 1967 in which he asks the tobacco lobbyists for another \$1,500 to keep writing. In mid-1968, the cigarette-affiliated Tiderock Corporation received this letter from Huff's publisher, Macmillan:

I must confess that we were not aware, except in connection with a grant for the preparation of the manuscript, that Mr. Huff had any past connections with the tobacco industry. Perhaps, as you suggest, it might be best for us to have a description of these connections, at this point.

Since we are anxious to complete our arrangements for the publication of Mr. Huff's work as soon as possible, it would be most helpful if your office could get the information to us very quickly. We'll be happy to have a messenger pick it up.

Sincerely,



William Mishkin  
Assistant Director  
Special Projects Division

But publication “as soon as possible” never seems to have occurred. What happened?

William Kloepfer, vice president for public relations for the Tobacco Institute, wrote of the manuscript, “Frankly, this mass of verbiage needs drastic editing before it will directly address itself to the needs of our industry.” After glancing at a couple of sections from a draft that was included in the documents library, I’d have to agree. “Mass of verbiage” is a pretty good description of Huff’s book.

In retrospect, I think Huff dodged a bullet. If *How to Lie with Smoking Statistics* had come out, I expect it would have destroyed his reputation—remember, this was 1969, five years after the surgeon general’s report—and taken a big bite out of the later sales and reputation of his 1954 bestseller.

How sincere was Huff? Did he tank his book for strategic reasons? I have no reason to believe Huff was lying or intentionally deceiving in his testimony. He may well have simply been misleading himself in analogizing research on the effects of smoking to silly things like studies of storks and babies. And if he was sincere in his views, I can hardly fault him for collecting some money for his efforts.

On the other hand, an internal document from December 1965 (<http://legacy.library.ucsf.edu/tid/vvr87e00>) makes me think Huff may have seen his role as producing talking points in support of a pre-determined conclusion. I guess we’ll never know if he really wanted to publish *How to Lie with Smoking Statistics*. Maybe he intentionally sabotaged it because he sensed it would ruin his reputation, whereas it was possible for him to keep the consulting and testimony under the radar.

## Part 2: The Statistician

Eminent statistics researcher (and, many years ago, my PhD advisor) Donald Rubin published a *Statistical Methods in Medical Research* article in 2002, “The Ethics of Consulting for the Tobacco Industry,” defending his work as an expert in court to dispute claims of legal liability of cigarette manufacturers, for which he was paid more than \$2 million over several years. Rubin’s article begins as follows:

This article describes how and why I [Rubin] became involved in consulting for the tobacco industry. I briefly discuss the four relatively distinct statistical topics that were the primary focus of my work, all of which have been central to my published academic research for over three decades: missing data; causal inference; adjustment for covariates in observational studies; and meta-analysis. To me [Rubin], it is entirely appropriate to present the application of this academic work in a legal setting.

I respect what Rubin is saying here—I don’t think he’d do this sort of consulting without thinking it through. I’d also like to highlight two complications.

At one point, Rubin writes, “When I was first contacted by a tobacco lawyer, I was very reluctant to consult for them, for the standard ‘politically correct’ reasons . . .” I think this is a bit glib. The phrase “political correctness” typically refers to attempts to restrict speech or ideology that is deemed offensive. Cigarette companies, on the other hand, actually make cigarettes, which actually do give people cancer. Now, I’m not saying it’s immoral to work for tobacco companies, or to supply cigarettes to people who want them, or even that it’s immoral to advertise cigarettes with the goal of inducing people to start smoking—I have no particular authority or inclination to set myself up as some sort of moral arbiter on this issue—but to dismiss such concerns as “political correctness” minimizes the issues here, I think.

Later in his article, Rubin presents the ethical dilemma of whether to give testimony that is scientifically valid but supports cigarette companies. Much of Rubin’s work in this job was spent discrediting flawed statistical analyses performed by the other side, and he makes a convincing case that, in his analysis, the facts did not support the claims made by his opponents. The topics that Rubin addressed were in the area of economics and causal inference; he did not address the health effects of smoking in his testimony. I would tend to accept Rubin’s reasoning that, once he had studied the issue, it was ethical for him to call the science as he sees it, even if that means he is supporting tobacco companies in a lawsuit. (If I had close personal experience with lung cancer victims—or with tobacco farmers—this would probably affect my views on this, but that’s another story.) Rubin also points out that flawed causal reasoning, once accepted in court, could have future malign effects; this is another motivation for him to argue on behalf of the best possible statistical analysis without regard to context. However, there is another pair of choice points that Rubin did not discuss in his article: his decision to work on the problem and then to take the cigarette manufacturers’ side of the case. Tobacco companies surely had their share of shoddy analyses that could be eviscerated by a world-class statistician.

Conditional on accepting the job as the cigarette companies’ statistician, I agree with Rubin that professional ethics demand that he give his best approximation to the truth. What is the best approximation? That’s a judgment call. As a statistician, I always try to use the simplest possible method that will do the job. Unfortunately, the simplest approach that does what is necessary is typically more complicated than anything I know how to do. Hence the need for research effort, even in what might seem like straightforward applied problems, and also hence the attraction of such consulting projects for a gifted statistician such as Rubin. It’s not just the money. This was a problem of which people really cared about the answer, which motivated the statistician to push harder to obtain and justify reasonable inferences.

### Part 3: The Industry

After hearing the fascinating, but nearly unknown, story of Huff's unpublished tract, I was motivated to read Proctor's history of cigarette manufacturers, their triumphant moves into mass production and marketing, and their struggles as they and the general public realized their product is both deadly and addictive.

My first stop upon receiving the book was the index, in particular the entry for Rubin, Donald B. I followed the reference to pages 440–442 and found the description of his activities to be accurate, neither diminished nor overstated, to the best of my knowledge.

Rubin is the second-most-famous statistician to have been paid by the cigarette industry (the most famous being R. A. Fisher, who at the end of his life stubbornly argued against the now universally accepted claim that cigarette smoking causes cancer). Here's a partial list of well-known statisticians who have worked for the cigarette industry:

Herbert Solomon, Stanford University

Richard Tweedie, Bond University

Arnold Zellner, University of Chicago

Paul Switzer, Stanford University

James Heckman, University of Chicago

Joseph Fleiss, Columbia University

Nathan Mantel, George Washington University

Joseph Berkson, Mayo Clinic

Much of the cancer-denial work was done after the 1964 surgeon general's report. For example, Proctor reports, "The statistician George L. Saiger from Columbia University received [Council for Tobacco Research] Special Project funds 'to seek to reduce the correlation of smoking and diseases by introduction of additional variables'; he also was paid \$10,873 in 1966 to testify before Congress, denying the cigarette-cancer link."

And here's a name that is famous to statisticians:

Ingram Olkin, chair of Stanford's department of statistics, received \$12,000 to do a similar job (SP-82) on the Framingham Heart Study . . . Lorillard's chief of research okayed Olkin's contract, commenting that he was to be funded using "considerations other than practical scientific merit."

Ouch. I bet that one didn't make it into the Stanford alumni magazine.

As late as 1974, a cigarette-company-funded pharmacologist published an article in *Executive Health* titled, "The Case Against Tobacco Is Not Closed: Why Smoking May Not Be 'Dangerous to Your Health!'"

This does not fit well with cigarette lobbyists' claims that everybody knew all along that cigarettes are

dangerous. People used to call them "cancer sticks," and so on. As Proctor demonstrates, surveys over the decades have found a lot of uncertainty about the health risks of cigarettes—and the cigarette companies were doing their best to prolong this uncertainty.

One thing I learned from Proctor's book was the distinction between tobacco and cigarettes. Tobacco is bad for you, but cigarettes concentrate the hazards through two key factors: mass production and how the tobacco is processed. What's the difference? Two key factors: mass production and how the tobacco is processed. Mass production means higher doses are more convenient and affordable (not such a good thing if you're addicted to a product that causes cancer). The part I didn't know about, before reading this book, is that the physical/chemical treatment (in particular, something called "flue-curing") makes cigarettes much less irritating to the throat, so that a smoker can more easily inhale and get those carcinogens directly into the lungs.

Thus, a world in which people grew tobacco in their backyards and rolled their own cigars would cut out lots and lots of smoking morbidity and mortality.

Proctor also notes the divergent interests of two groups that are often conflated: cigarette companies and smokers. According to surveys cited by Proctor, most smokers (in the United States, at least) want to quit. I'm assuming most cigarette companies don't want this to happen. This is all well known, but it sometimes gets lost in discussions of the "war on smokers," etc.

Proctor also reports some amazing court testimony from Kenneth Ludmerer, a professor of history and medicine at Washington University in St. Louis:

Q: Doctor, is it your opinion that cigarette smoking contributes to the development of lung cancer in human beings?

A: I have no opinion on that.

This was a professor at one of the nation's leading medical universities, testifying in 2002. I wonder if he still has no opinion on whether "cigarette smoking contributes to the development of lung cancer in human beings." Maybe there have been some important research developments since 2002 that have convinced him.

As a professor, it is natural for me to become particularly indignant about the offenses of my academic colleagues and to prefer Rubin's open defense of his consulting to what seems to me as the slippery evasions of a medical doctor who claims to have no opinion on whether smoking contributes to the development of lung cancer.

To his credit, though, Proctor doesn't lose track of the larger story, which is the century-long transition of tobacco-smoking to mass-produced lethality. In the first part of the twentieth century, cigarette companies conducted research into smoking and cancer with the hope of developing a safe cigarette. After all, they had no motivation to kill their customers. It eventually



became clear that the safe cigarette wasn't going to happen. Here's Proctor:

The year 1953 marks a turning point of sorts ... Smoking was charged with causing cancer, and popular media were reporting on the facts. What was the industry to do? ...

Philip Morris Vice President George Weissman, in March 1954, announced that his company would 'stop business tomorrow' if 'we had any thought or knowledge that in any way we were selling a product harmful to consumers.' James C. Bowling, the public relations guru and Philip Morris vice president, asserted in a 1972 interview, 'If our product is harmful ... we'll stop making it.' Then again in 1997, the same company's CEO and chair, Geoffrey Bible, was asked (under oath) what he would do with his company if cigarettes were ever established as a cause of cancer. Bible gave this answer: 'I'd probably ... shut it down instantly to get a better hold on things.' The other manufacturers made similar assurances.

Lorillard's president, Curtis Judge, is quoted in company documents as saying, 'If it were proven that cigarette smoking caused cancer, cigarettes should not be marketed,' and he would 'quit his employment.' R.J. Reynolds president, Gerald H. Long asserted in a 1986 interview that if he ever 'saw or thought there was any evidence whatsoever that conclusively proved that, in some way, tobacco was harmful to people, and I believed it in my heart and my soul, then I would get out of the business.'

This all makes a lot of sense, but it's a lot different from what I was hearing when the cigarette industry was on trial several years ago. Then, the story was that everyone had known forever that smoking caused cancer, and that cigarette manufacturers were performing the useful service of supplying a consumer good that many people wanted. It's interesting to see, at least in public, cigarette executives taking a much more direct position that they did not want to be in the position of giving people cancer: "If our product is harmful ... we'll stop making it."

## Part 4: Statistics

Statistics is central to the controversies over smoking and health, from the epidemiological evidence tying cigarette smoking to mortality from cancer, heart disease, and other sources, to analysis of animal experiments, to the presentation of data (one trick of tobacco publicists was to present trends in several causes of death, excluding lung cancer), to the debates discussed by Rubin of the economic costs to society of smoking-related illnesses.

And some of the statistical questions are fascinating and far from trivial. The rules promulgated by Austin Bradford Hill for assessing causation in observational studies—inspired in large part by studies of smoking and cancer—remain influential and have recently been re-expressed and extended by Paul Rosenbaum into a statistical framework of causal inference. Social studies of cigarette behavior have inherent multilevel structures: at the individual level, many anti-smoking programs have been found to be ineffective (which is one of Rubin's reasons for arguing that, even had tobacco companies warned their consumers of cancer risks, this would have had little effect on smoking rates). At the same time, however, smoking habits vary dramatically between states, countries, and demographic groups, suggesting that, paradoxically, these drug-taking behaviors are more easily altered in the aggregate than individually.

To return to the ethical questions, it is inevitable that statisticians will work for all sorts of employers. Some of my research is funded by the U.S. Department of Defense, an organization many would argue is more dangerous than any tobacco company. When working for a controversial organization, I believe we have an ethical obligation to uphold our best standards of statistical work, to speak honestly, and to be open about our financial interests, all of which Rubin did. An ethical argument could be made on the indirect effects of testimony (by poking holes in the anti-cigarette arguments rather than the other side's, the effect could ultimately be to keep the cigarette industry alive longer and thus lead to more deaths), but then we are getting to broader ethical questions and leaving the realm of statistical ethics, which is the subject of this column.

Was Huff being ethical in brainstorming with cigarette lobbyists about how best to mock the cancer threat? It's hard for me to say. Huff was acting not as a statistician but as a journalist, or, more precisely, a public relations operative. The Public Relations Society of America does, in fact, have an ethics guide, the first principle of which is to "protect and advance the free flow of accurate and truthful information." In this case, it is reasonable to suppose that Huff, in 1968, (unlike Ludmerer in 2002) had no capacity for judging the evidence on smoking and cancer and was willing to believe what his clients told him. Huff's tobacco work could well be considered less than admirable (as evidenced by the fact that even his admirers seem to have been unaware of it) while not violating the standards of professional ethics. ■

## About the Author

**Andrew Gelman** is a professor of statistics and political science and director of the applied statistics center at Columbia University. He has received many awards, including the Outstanding Statistical Application Award from the American Statistical Association and the award for best article published in the *American Political Science Review*. He has coauthored many books; his most recent is *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*.