J Quant Criminol (2010) 26:425–428
DOI 10.1007/s10940-010-9112-8

BRIEF ESSAY

Picturing JQC's Future

Michael D. Maltz

Published online: 22 September 2010
© Springer Science+Business Media, LLC 2010

Like Al Blumstein, I came to criminology via operations research (OR) and, earlier, via electrical engineering, when I was engaged in designing a radio system for the Boston Police Department (Maltz and Waldron 1968; Maltz 1970). I subsequently became an OR analyst for the National Institute of Law Enforcement and Criminal Justice (now NIJ) in 1969. After 3 years of seasoning, I started teaching criminal justice at the University of Illinois at Chicago Circle (now UIC without the Circle[1]). There I taught *inter alia* undergraduate and graduate statistics and research methods. Having never taken them, I had to learn them on the fly. But like most engineers, I had a strong background in dealing with data, so it was not too problematic.[2]

But I brought different ways of looking at data (with an emphasis on the word *look*). Engineers and most hard scientists are taught to plot the data as they collect them. In physics we added weights to a spring and plotted the spring's length against the applied weight; we dropped an object and plotted the time it takes to fall different distances. It is through such exercises that we were able to demonstrate the linear relationship in the first example and the quadratic relationship in the second.

In becoming a criminologist, however, I found that I had to deal with different types of data, for the most part based on random samples. The National Crime Survey (NCS), and its successor, the National Crime Victimization Survey (NVCS), were based on samples, as were a number of other important data sets. I set out to learn the methods of dealing with

[1] I guess I've been around long enough to see a lot of name changes. The "circle" stood for the confluence of three Chicago expressways, the Kennedy, the Eisenhower, and the Dan Ryan (no, he wasn't a president you hadn't heard about). We used to joke that it was the only university named after a traffic intersection.

[2] One of my first students at UIC was John Laub, I think in my first year of teaching, so I must not have been too terrible a teacher.

M. D. Maltz (✉)
Ohio State University, Columbus, OH, USA
e-mail: mdm@sociology.osu.edu

M. D. Maltz
University of Illinois at Chicago, Chicago, Illinois, USA

sampled data, including the use of confidence intervals and statistical significance in making decisions about the value of findings.

In fact, the first paper I wrote that included an analysis of confidence intervals was one that estimated the gain and loss in state revenue due to cigarette smuggling (Maltz 1976, 1981). It resulted in my getting a nice letter from someone I had never heard of at that time, W. Edwards Deming; he sent me some of his papers, even though "I doubt if they will be of any interest to you." Of course, as I became more interested in statistical analysis I began to consult those papers! One in particular, "On probability as a basis for action" (Deming 1975), stands the test of time.[3]

As the years passed I noticed that more agencies were computerizing their data, and data began to flow in ever-increasing streams. Statistical significance was no longer a good measure of the value of a finding. The turning point for me was my review of a report on recidivism submitted to the National Institute of Justice. The sample the researchers had to work with was so large that every finding was significant, which made their tests meaningless. To deal with this problem, they decided to take a ten percent sample of the cases and analyze the sample. In that way not all of the findings were significant, so they could produce a report that was less confusing to them.

Consider what they did from another standpoint: in order to produce what they felt was a good analysis, they threw out 90 percent of the data! As someone who was and is a data maven, I felt that this bordered on the criminal. From that point on, I began to turn my thoughts to how to promote more useful methods, methods that not only didn't rely on sampling methods, but rather were tailored to deal with the increasing quantities of data that we were experiencing. And among the most useful ones I found were those laid out in John Tukey (1977) book, *Exploratory Data Analysis* (EDA), which developed a number of graphical techniques. He distinguished EDA from *confirmatory* data analysis and noted that (p. vi).

> "once upon a time, statisticians only explored. Then they learned to confirm exactly – to confirm a few things exactly, each under very specific circumstances. As they emphasized exact confirmation, their techniques inevitably became less flexible. The connection of the most used techniques with past insights was weakened. Anything to which a confirmatory procedure was not explicitly attached was decried as 'mere descriptive statistics', no matter how much we had learned from it."

This resonated with me, in that so many social scientists at the time put down actually looking at the data as "data dredging" rather than as a natural step in exploring the characteristics of the data. In focusing on exploratory techniques, however, Tukey was constrained by the computing environment at the time, which was horse-and-buggy compared to our current capability. He therefore relied on simple computations and manual graphs, since computers at that time were unable to display graphs. His "metamethods," however, are still applicable. LOOK at the data; PLOT them to see if any relationships jump out; TRANSFORM the data (logs, powers, exponentials, etc.) to see if they clarify the plots; DISAGGREGATE the data to see if subsets have their own patterns; TRY different relationships; and REPEAT until satisfied. The drudgery that used to be associated with this sequence has all but disappeared with the advent of new graphically-based software attached to database management systems; however, they had not been tried to

[3] One quote from that paper: "Little advancement in the teaching of statistics is possible, and little hope for statistical methods to be useful in the frightful problems that face man today, until the literature and classroom be rid of terms so deadening to scientific enquiry as null hypothesis [and] level of significance for comparison of treatments.".

any great extent when I assumed the editorship in 1997, and I was eager to "lead the journal into new methodological areas" (Maltz 1997: 93).[4]

I was young(er) then and had ambitious plans, and thought that all I needed to do was describe what I wanted to do, and the submissions would follow. The direction I was proposing was to some extent laid out in an earlier article in the *Journal of Research in Crime and Delinquency* (Maltz 1994). Specifically, I wanted the journal to promote not just the development and use of new statistical algorithms, but the development of visual techniques. I was then in the middle of a 5-year visiting fellowship at the Bureau of Justice Statistics ("Development of Graphical and Geographical Methods"), with a book and a number of articles and chapters describing how data visualization could be used in analysis (Maltz 1981, 1984/2001, 1994, 1996; Maltz and Pollock 1980; Maltz et al. 1991/2000).

I had, however, underestimated the extent of inertia in the academic enterprise. Most social science methodologists are trained by other social scientists who came of age in the SAS/SPSS/Stata/Systat era, during which time the methodologies that were taught were largely those available in these programs, and were dominated by confirmatory (as opposed to exploratory) techniques. The developers of these software packages made it easy to perform confirmatory data analyses: with the click of a few buttons, a researcher could transform a data set, apply a procedure to it, come up with a result, if necessary tweak the data or the model, apply additional procedures, and continue to play with the data and methods until a (statistically significant) finding emerged.

This, of course, is an oversimplification, but unfortunately still rings true. What I had hoped for was to broaden the armamentarium of criminology and criminal justice researchers to include Tukey's exploratory data analysis, as implemented by Tufte (1983) and Cleveland (1993, 1994), among others.

While EDA methods show how to get insight from raw data, I also had hoped that graphical techniques would also be used to describe the extent of uncertainty in statistical parameters. Toward that end Marianne Zawitz and I prepared a report (Maltz and Zawitz 1998) for the Bureau of Justice Statistics (BJS) showing how the National Crime Victimization Survey data could be depicted to include their confidence intervals, and therefore provide a more intuitive (to us, at least) understanding of NCVS findings. [BJS used our graphical techniques in their annual reports on the NCVS subsequently, but this was cut back in later years, and discontinued completely more recently. Zawitz's recent retirement from BJS may have been the reason for this, since she prepared the graphs for the NCVS reports.]

Despite these setbacks, I foresee the use of graphical techniques increasing in JQC in the future. Aside from their use (and dominance) in geographical criminology, they will augment those who use graphical techniques to ferret out patterns in the huge data sets that become available. That is, I expect there to be less emphasis on basing statistical methods on sampled data and more on data mining, less on deductive methods and more in inductive methods. I expect to see JQC and other journals migrate to the Internet, permitting others direct access to articles and to data sets used or referenced in the article. I expect to see animated graphics, similar to those produced by Hans Rosling's Gapminder (2010), providing readers with insights into variables' relationships heretofore unavailable.[5] My hope is that JQC remain at the forefront of these advances.

[4] The first paper following this editorial was an example of EDA, in which Brian Forst and Jim Lynch (1997: 97–119) draw isoquant maps to "reveal patterns that are often obscured" when using conventional methods.

[5] Google recently started producing animated bubble charts similar to Gapminder—see http://www.google.com/publicdata/home (accessed March 9, 2010).

References

Cleveland WS (1993) Visualizing data. Hobart Press, Summit, New Jersey
Cleveland WS (1994) The elements of graphing data. Hobart Press, Summit, New Jersey
Deming WE (1975) On probability as a basis for action. The American Statistician 29(4):146–152
Forst B, Lynch JP (1997) The decomposition and graphical analysis of crime and sanctions data: a cross-national application. J Quant Criminol 13(2):97–119
Maltz MD (1970) Police communication–a case history. In: Mensch AMJ (ed) The survival of communication networks. NATO, Brussels, Belgium
Maltz MD (1976) On the estimation of smuggling in a 'gray market' commodity. Oper Res 34(6): 1156–1163
Maltz MD (1981) Transportation modeling in analyzing an economic crime. In: James AFox (ed) Methods in quantitative criminology. Academic Press, New York, NY, pp 77–97
Maltz MD (1984/2001) Recidivism. Academic Press, Orlando, Florida, 1984. Internet version available at http://sociology.osu.edu/mdm/Recidivism.pdf
Maltz MD (1994) Deviating from the mean: the declining significance of significance. J Res Crime Delinq 31(4):434–463
Maltz MD (1996) From Poisson to the present: applying operations research to problems of crime and justice. J Quant Criminol 12(1):3–61
Maltz MD (1997) Editor's introduction. J Quant Criminol 13(2):93–96
Maltz MD (1998a) Visualizing homicide: a research note. J Quant Criminol 14(4):397–410
Maltz MD (1998b) Which homicides decreased? Why? J Crim Law Crim 88(4):1479–1486
Maltz MD, Waldron S (1968) A supervisory communications system for the Boston Police Department. In: Cohn SI (ed) Law enforcement science and technology II. IITRI, Chicago, Illinois
Maltz MD, Pollock SM (1980) Analyzing suspected collusion among bidders. In: Geis G, Stotland E (eds) White-collar crime: theory and practice. Sage, Beverly Hills, California, pp 174–198
Maltz MD, Zawitz MW (1998) Displaying violent crime trends using estimates from the National Crime Survey, Report No. NCJ-167881, Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice, Washington, DC, June, 1998, 10 pp. Available at http://www.ojp.usdoj.gov/bjs/abstract/dvctue.htm
Maltz MD, Gordon AC, Friedman W (1991/2000) Mapping crime in its community setting: event geography analysis. Springer-Verlag, New York. Internet version available at http://sociology.osu.edu/mdm/mappingcrime.pdf
Rosling H (2010) Gapminder: unveiling the beauty of statistics for a fact based world view. http://www.gapminder.org/. Accessed 5 Feb 2010
Tufte ER (1983) The visual display of quantitative information. Graphics Press, Cheshire, Connecticut
Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading, Massachusetts