Comment

Stephen E. Fienberg

To comment on this paper is a somewhat daunting task for me. I usually find Dempster's writing provocative and thoughtful, and the present paper is no exception. In addition, I agree completely with Dempster's subjective perspective and his views on causality. In fact, my views on these have been strongly influenced by his classes and writings on these topics. Moreover, the contents of the paper appear to be technically correct, at least from the perspective of mathematical statistics. Yet, having read the paper now several times, I have concluded that the mathematical formalism and the discussion fail to elucidate the challenge that the issue of employment discrimination presents for statistical science. My difficulties come not from Dempster's manipulation of regression models, both direct and reverse, but rather from the language used to describe and interpret those models and in the attempt to explain how such models relate to the issues of discrimination in the workplace and the role of statisticians as expert witnesses in employment discrimination litigation.

My remarks reflect ongoing concern regarding the proper role of statistics and regression-like models in the social sciences (e.g., see the discussion in my exchange with Freedman on the topic—Freedman, 1985a, b; Fienberg, 1985), experience as a statistical expert in several legal cases of diverse natures involving employment discrimination, and continuing interest in the actual use of statistical methods in litigation and the influence they have on the resolution of legal issues (e.g., see Fienberg, 1988). Those of us with interests in the legal arena continue to look with horror on the ways in which statistical ideas and methods are misused over and over again by expert witnesses (often not statisticians or even those trained in statistics) or misinterpreted by judges and juries. I, too, agree with Dempster and Pratt (1986) that "we must keep striving toward sensible modes of using statistics in legal and public arenas" because I have seen the alternatives! I am not sure that Dempster's analysis in the present paper, as interesting as it is technically, does much to move us forward toward our mutual goal.

Stephen E. Fienberg is Maurice Falk Professor of Statistics and Social Science and Dean of the College of Humanities and Social Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213.

My principal difficulties with Dempster's paper stem from his account of the econometric model for employment discrimination. I simply don't think that it captures the mechanisms and processes used by employers to set salaries, and I certainly don't accept the role of the n true worth measures (Y_i^{**}) in defining *fairness*. Once you buy into the linear model framework that follows from Dempster's causal mechanism and the calculation of $Y^* = E(Y^{**} \mid G, X^*)$, everything else follows but the interpretation. Instead, I would argue for a framework which allowed the statistician to focus on employment process decisions and to learn about the input to those decisions. For related comments see Michelson (1986).

A brief story will illustrate the nature of my concern. In the late 1970s, a class action lawsuit was brought against a large southern employer, alleging systematic employment discrimination against women involving hiring, compensation and promotion. When the case went to trial, the centerpiece of the plaintiffs' argument was the testimony of a statistical expert who carried out multiple regressions galore, using a model much like Dempster's equation (1). She concluded that the coefficient for gender was significantly different from zero and proceeded to estimate the damages, using the causal model argument Dempster outlines and rightly deplores. To rebut this evidence the defendant put on the stand a statistical witness who did just what Dempster suggested; he added a term $X'\beta'$ to the models of the plaintiffs' expert and noted the extent to which the estimated gender coefficient changed. The judge, in his written opinion, stated that neither of the experts' regressions had anything to do with the realities of the case, but found for the plaintiffs nonetheless. A year or so later someone shared with me the trial transcript and some of the experts' statistical exhibits. As a statistician I was appalled by the naive analysis and the uncritical acceptance of the regression framework and its causal interpretation by both experts. Shortly thereafter, I was lecturing in a multiday seminar on the topic of employment discrimination and, at one of the breaks, I struck up a conversation with one of the participants. He turned out to be an attorney for the company that was the defendant in this case. I immediately told him what I thought about the use of statistical evidence by the two witnesses and asked: "Why did the company's lawyers allow their expert to present such mindless regression analyses in response to the equally mindless ones

of the plaintiffs' expert?" He said to me: "You don't understand. If the plaintiffs' expert hadn't been busy running multiple regressions she might have taken a closer look at the employee manual which describes what in essence is a two-tiered job system. Men are channelled into one tier and women into the other. After that, virtually all employment decisions follow as a matter of course. When our expert responded by running his own regressions, the lawyers were quite pleased. They believed that the outcome would have been far worse if he had explained to the court what we really do because then the judge could easily have concluded that our system was discriminatory on its face."

Within Dempster's framework, I had special difficulty in understanding the distinction he attempts to draw between judgmental discrimination and prejudicial discrimination. For me, attributing judgmental discrimination to "a presumed honest attempt to assess productivity" is ignoring the realities of the legal meaning of discrimination and the judicial injunction that statisticians cannot use intrinsically tainted carriers of discrimination as predictors in their statistical models. It is all well and good for Dempster to say that his definition of fairness implies that "there is no restriction at all on the variables admitted to X^*," but it won't do him much good if he attempts to take his framework into the courtroom. This is the problem I alluded to at the beginning of this comment. When statisticians use labels with nonstatistical, value-laden meanings to interpret coefficients and variables in an abstract statistical model, they cannot hope to advance statistical science. Nor can they expect agreement on the interpretation of their statistical efforts in adversarial settings.

ACKNOWLEDGMENT

This work was facilitated in part by Grant SES-87-01606 from the National Science Foundation.

ADDITIONAL REFERENCES

FIENBERG, S. E. (1985). Comments on and reactions to Freedman, statistics and the scientific method. In *Cohort Analysis in Social Research* (W. M. Mason and S. E. Fienberg, eds.) 371–383. Springer, New York.

FIENBERG, S. E., ed. (1988). *The Evolving Role of Statistical Assessments as Evidence in the Courts*. Springer, New York. To appear.

FREEDMAN, D. A. (1985a). Statistics and the scientific method. In *Cohort Analysis in Social Research* (W. M. Mason and S. E. Fienberg, eds.) 343–366. Springer, New York.

FREEDMAN, D. A. (1985b). A rejoinder to Fienberg's comments. In *Cohort Analysis in Social Research* (W. M. Mason and S. E. Fienberg, eds.) 385–390. Springer, New York.

MICHELSON, S. (1986). Comment on "Regression analyses in employment discrimination cases" by D. A. Conway and H. V. Roberts. In *Statistics and the Law* (M. H. DeGroot, S. E. Fienberg and J. B. Kadane, eds.) 169–181. Wiley, New York.

Rejoinder

Arthur P. Dempster

1. FRANKLIN FISHER

Much of Franklin Fisher's commentary consists of adversarial argumentation of a sort often heard in courtrooms. In my paper, I mainly kept discussion of active legal processes in the background, because the issues I was discussing were intended to be primarily scientific. But I accept that it is fair tactics on his part, given that our relationship apparently continues to be adversarial in the scientific realm, to bring out that my practical experience was primarily in advising counsel and testifying on behalf of defendants (i.e., employers), while he served on behalf of plaintiffs (i.e., in some cases one or more employees who believed themselves to be victims of discrimination, or in other cases the government acting on behalf of a protected class of employees whether or not grievances had been registered).

That we chose sides as we did is presumably not a chance result. For my part, I believe that the explanation has nothing to do with a predilection to find for one side or the other. Rather, my preference resulted from a conviction that the statistical strategies typically pursued by plaintiffs in employment discrimination cases were serious flawed, as I continue to believe. No doubt Fisher can offer a parallel explanation for his choice of side. But the symmetry ends there, for he evidently feels that the validity of direct regression methods is such that plaintiffs' cases are often proved by statistical arguments, whereas my expert view of the epistemic deficiencies of many plaintiffs' experts' statistical arguments suggests that no statistically based judgments should be reached until the defects in the arguments are repaired. The repairs will be difficult and demanding in terms of commitment of professional resources, because they