

Comment

Harry V. Roberts

We appear to have passed through the period of most intense litigation of legal cases alleging employment discrimination in compensation or promotion by race or sex. The highest frequency of academic papers on the statistics of age discrimination appears also to be past.

Arthur Dempster's paper therefore comes at a good time for an overview of the statistical issues raised by race and sex cases. In my view, Dempster's overview is fascinating, authoritative and close to comprehensive on the foundations of discrimination study with statistical tools. The paper not only deals with discrimination studies in particular, but illuminates the vexing general problem of causal inference from observational data where explicit randomization has not or even cannot be applied.

The paper warns also about the dangers in the belief that sufficiently sophisticated models can completely resolve differences over causal interpretation: "statistical analysis on its own rarely offers complete solutions to externally specified problems." Dempster stresses the need to "look outside the data for evidence bearing on the missing information."

I find nothing to criticize in Dempster's general development or in his model based on Bayesian principles. At the same time, I feel that more attention to certain concrete details would help to deal more specifically with important questions that have been debated by statisticians inside courtrooms and in the pages of scholarly publications. I shall therefore attempt to fill in details that I believe to be important.

Part of my comment will be directed to an amplification of Dempster's comment about "looking outside the data." Looking outside the data entails examination of the employment process that produces the data. Such examination reveals specific activities such as search for job candidates, selection from applicant pools, initial placement and salary determination, salary advancement and job promotion and even reduction of force, which has come to be important in the statistical analysis of age discrimination cases. In looking at these activities, one is led to flesh out the econometric model of Section 3. One is led also to consider legal questions. One is made more aware of the limitations of available data, including both job

qualifications and employment history. For example, good job performance measures are seldom available and information is often lacking about job applicants who were not hired.

But one can also look *within the data*: mundane data analysis can provide clues about causal analysis. It was through data analysis in a series of discrimination cases from 1974–1985 that I gradually acquired a better understanding of what happens in the employment process that is relevant to a statistical study of the process. That understanding led to modifications and extensions of the methodology of earlier studies, including first the introduction of reverse regression and later the modification of the way reverse regression is applied. Consideration of problems that I encountered in data analysis leads to a fuller understanding of the issues so well presented by Dempster. (This approach also gives an autobiographical tone to my comment that I regret but cannot avoid without blunting the points to be made. I regret also the necessarily extensive self-citations in the references for this comment.)

During the process of data analysis, not only my statistical methodology but my views of the important causal issues evolved. The evolution is reflected in the series of papers and reports listed under "Additional References." (The papers from 1983 on were jointly authored with Delores A. Conway. I cannot overstate the importance of Professor Conway's contributions to my own thinking, but because we are preparing separate comments on Dempster's paper that we have not had an opportunity to discuss with each other, I hasten to say that she may not fully agree with all aspects of this comment.)

ADVANCEMENT STUDIES

In 1974–1975 I worked with Mary Townsend Kimpton on a salary discrimination case involving United Airlines (Field, Kimpton, McGee and Roberts, 1978). Study of salary discrimination at hire was constrained by lack of relevant information on job qualifications. Hence, salary advancement after hire seemed to be the only aspect of the case amenable to regression analysis. Hence, our regression study conditioned on starting salaries and on the available (limited) information about job qualifications. I came later to refer to such a study as an "advancement study."

It is interesting to observe that at that time, the plaintiff presented only simple tabular comparisons of

Harry V. Roberts is Professor of Statistics, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637.

male/female and white/minority salary and job levels, with no attempt to condition on covariates.

PLACEMENT STUDIES

In 1977 I began work on the Harris Bank case. A little earlier I had started work on a case for another large bank that never ultimately came to trial, although the statistical work was comparable in scope and intensity to that on the Harris case. In both cases, data on employee qualifications were fuller than in the United Airlines case, although still quite limited by comparison with what one might wish. This greater availability of data made it feasible to supplement an advancement study by a regression analysis of salary at hire, which I referred to as a "placement study." I argued that the division of a discrimination study into two stages, placement and advancement, is advantageous because it provides greater specificity in diagnosis and, if the need is indicated, in remedy.

At both banks it was possible to classify entering jobs into a small number of relatively homogeneous subgroups. Salary regression analyses of employees hired within a limited time interval could be conditioned on these entering subgroups. I did salary regressions both with and without subgroup conditioning, so that the data could speak more clearly about what was happening.

REVERSE REGRESSION

Of course, salary regressions conditioned on entering job subgroups could not detect possible discrimination in hiring into these subgroups, a form of discrimination sometimes called "shunting" because females or minorities could be directed or shunted into jobs for which they were over qualified. (In the two bank cases there was little or no information about applicant pools and the legal charges were confined to salary discrimination, so no direct study of shunting was possible or appropriate.)

In the process of preliminary placement studies at both banks in 1978, I noticed that mean years of education (the major qualifications variable in salary regressions) were nearly the same for males and females within each entering job subgroup. It seemed to me that this finding of data analysis was contrary to what I would have expected if shunting had been practiced by the banks. A few weeks before the start of the first Administrative Hearing in the Harris Case in August of 1979, it occurred to me that the idea could be extended from entering job groups to groups of employees with closely similar salaries. That is, one could study male and female employees with the same salaries and compare their education or other job qualifications. In regression terms, one could regress

education, for example, on salary and an indicator variable for sex. I called this approach "reverse regression." The idea was of course not novel, even in the area of discrimination studies, where I later learned that Birnbaum (1979) had applied the same general concept.

For procedural reasons unrelated to my work, my Harris studies were not admitted into evidence in the Harris Administrative Hearing of 1979, and my role in the hearing was confined to rebuttal. After the hearing, however, the Harris Bank permitted me to prepare and circulate a summary of placement and advancement regressions (Roberts, 1979). In that summary, I presented direct as well as reverse regressions and offered reasons for preferring the reverse regressions. In retrospect, three rather different reasons were in my mind and were expressed with varying degrees of clarity in my report and in my rebuttal testimony during the hearing.

The primary reason was based on the idea of a fallible covariate. That is, job qualifications were very imperfectly measured whereas salary could be measured precisely, so the reverse regression might come closer to a correct causal picture. A paper by Goldberger preliminary to the one referenced in Dempster's paper convinced me that this reasoning was not really relevant to my regression studies of employment discrimination. Had qualification variables consisted of scores on aptitude or achievement tests, these would have constituted a fallible covariate. But qualifications such as years of education, company seniority, age and prior job experience would not so qualify.

A second reason, pointed out to me first by Dempster in August of 1979, was the idea of fairness, which Delores Conway and I subsequently developed systematically (Conway and Roberts, 1983). We pointed out that direct and reverse regressions correspond to two different senses of fairness, and that it is often impossible to have both. Fairness arguments do not point unambiguously either to direct or reverse regression, but they show how different ideas about fairness or equity imply one or the other statistical mode.

EQUAL PAY FOR EQUAL WORK

The third reason was an informal attempt at causal modeling, but it was not well expressed in 1979. Gradually I began to understand it more clearly. In so doing I was aided in part by the Goldberger-Dempster exchanges, but mainly by practical data analysis, reflection on the results of data analysis and, above all, the keen insight of Delores Conway. The data analysis came about thus. In the Ward's case discussed

in Roberts (1985a), the initial complaint required the defense to prepare a study of comparative salary advancement for males and females starting in September of 1975. Under the Civil Rights Act of 1964, the company was not responsible for any male-female salary disparities that might have existed in September of 1975. As analysis proceeded, it became clear that there was approximate parity in salary advancement between males and females after September 1975.

At this point, the plaintiff's attorneys turned their attention away from the Civil Rights Act to the Equal Pay Act of 1963, which would have made the company responsible for existing male-female salary disparities. The attorneys thus sought to prove discrimination in the sense of unequal pay for essentially equal work.

HOMOGENEOUS JOB GROUPS

This shift in legal orientation led naturally to comparison of male and female salaries by regression analyses within subgroups of employees doing essentially equal work, that is, within what Delores Conway and I later called relatively "homogeneous job groups."

There are many ways of defining homogeneous job groups, none of them completely satisfactory. Essentially one is involved in a process of successive disaggregation, starting from the entire workforce at issue in a case, say all white collar workers at a given location. Disaggregation might begin by separate consideration of the two subgroups defined by all professional employees and all clerical employees, and then be extended to still finer subgroups within professional and within clerical employees that could be defined by company job classification schemes. Within any subgroup, direct regression can be viewed as an attempt to model the process by which an employer is presumed to "set" salaries, because salary determination is mainly carried on within relatively homogeneous subgroups of employees; the subgroups are noncompeting with respect to each other.

In the context of an equal pay case, therefore, the argument for direct regression seems natural and compelling, although of course the usual problems of omitted variables can cloud the interpretation. The major problem of implementation is the degree of disaggregation to be attempted, because more disaggregation means smaller individual sample sizes and less aggregation means more heterogeneity within subgroups. A summary of my attempts to deal with this problem in the Ward's case is contained in Roberts (1985a); a report of subsequent refinements is contained in Conway and Roberts (1987). Parallel analyses prepared for the second Administrative Hearing in the Harris Bank Case are summarized in Roberts (1985b).

Within the broader context of civil rights cases, however, many statisticians appear to be reluctant to condition salary regression analyses on homogeneous job groups because possible discrimination in initial selection into such job groups cannot be detected by any salary regression confined only to those actually selected. This position puzzles me because it appears to assert that an analysis should be avoided if it cannot answer all questions of importance in a case even if it can help to answer some of them.

Delores Conway and I recognized the need for a framework that would deal with the possibility of discrimination both in selection and in salary determination. An outline of that framework is contained in Conway and Roberts (1986b, 1987). Briefly, we begin by saying that if data are available on, say, an applicant pool for a given subgroup of jobs, a reasonable approach to study of selection discrimination is to use *direct* logistic regression of selection/nonselection on measured job qualifications and an indicator variable for sex. For study of salary regression, we proposed *direct* regression of salary on job qualifications and an indicator variable for sex.

Therefore, *so long as one conditions on appropriately homogeneous groups of employees or applicants*, whether one is studying quantitative rewards such as salaries or qualitative rewards such as selection, direct regression is seen as the tool of choice. Direct regression can claim to model employer behavior because selection and salary determination occur in the context of given jobs or closely related jobs. Thus, in a university, one would not mix faculty and clerical employees in a single direct regression, but a direct regression for faculty separately and clerical employees separately would be appropriate; and faculty members could be further subdivided by departments or groups of related departments. (Again, in all regressions, direct or reverse, the problems of causal interpretation in the face of omitted or unavailable variables must be faced.)

But there still are roles for reverse regression. For example, Conway and I (1986a, 1987) have argued that full data analysis entails viewing the data from different perspectives, and that reverse regression is one useful perspective. Further, we argued that a modified use of reverse regression can be used in study of selection discrimination as a partial compensation for the lack of detailed data on applicant pools. This modification is explained next.

MODIFIED REVERSE REGRESSION

There is an important role for a variant of reverse regression in the study of selection discrimination in the common circumstances in which data on applicant

pools are missing. This variant goes back to the simple reverse regressions of placement that I had done in 1978: one compares, say, male and female job qualifications within homogeneous job groups. The idea of detection of possible shunting now, as then, seems attractive, and I would describe such an analysis as a "shunting study."

When multiple job qualifications are available, the full multivariate comparison can be reduced to a univariate comparison by the same device I used in the salary reverse regressions of 1979. A univariate index of job qualifications is constructed from the fitted values, less the contribution of the sex indicator variable, of a salary regression for the same subgroup. This index is simply an estimate of how employers weight available job qualifications in "setting" salaries for those selected.

As pointed out by Arlene Ash in her discussion of Conway and Roberts (1986b), modified reverse regression cannot be conclusive, even granting the usual reservations of limited sample size and omitted variables: *Parity on a shunting study does not imply parity on the full logistic regression that would be possible if data on rejected applicants and their qualifications were available.* Nonetheless, the modified reverse regression can be useful. A finding of disparity in a shunting study would constitute evidence in support of discriminatory selection or placement; and conversely, a finding of parity can provide some support for nondiscriminatory selection. In either instance, even casual evidence about the comparative qualifications of males and females in candidate pools can be helpful in interpretation of the statistical analysis. Here is a nice illustration of Dempster's point: "statistical analysis on its own rarely offers complete solutions to externally specified problems."

DISAGGREGATION VERSUS AGGREGATION

The final disaggregated approach sketched above and set forth in detail in Conway and Roberts (1986b, 1987) was reached only gradually after more than 10 years of data analysis on perhaps a dozen cases involving extensive data sets (most of the cases were settled without coming to trial). It differs substantially from the more usual aggregate analyses, which are aggregated studies of the entire workforce defined by the legal scope of the case. In these analyses, company job structure and homogeneous job groupings usually have little or no role. (I do not oppose aggregated studies; they are part of the picture, and they are shown in Roberts (1985a, 1985b), where modified reverse regression was used.)

However, as explained in Conway and Roberts (1986b), in the absence of discrimination within homogeneous job groups, aggregated salary regressions would show apparent discrimination *so long as*

the mean job qualifications for the protected group are less than those for the nonprotected group. This latter condition prevails in almost all aggregated regression studies known to me, including my own. Disaggregation, by contrast, tends to reduce or even remove the appearance of discrimination, as illustrated in Roberts (1985a, 1985b).

I believe that the emphasis on aggregated versus disaggregated analysis is the most important explainer of differences in conclusions between expert statisticians in discrimination cases. The differences cannot be resolved by a decisive significance test or theoretical arguments about the merits of various formulations of an econometric model. They can be isolated by comparisons of aggregated and disaggregated analyses of the same data base.

JUDGMENTAL DISCRIMINATION AND PERFORMANCE EVALUATION

Finally, I want to consider briefly an unrelated point, Dempster's development of the concept of judgmental discrimination. I find his treatment thought provoking and useful. However, even if some kinds of judgmental discrimination were to be made legally defensible, good performance evaluations would seem to be a necessary prerequisite for a statistical defense for judgmental discrimination. Unfortunately, many rating systems appear to be largely judgmental and even capricious. Indeed the absence of defensible rating systems is a reason for the need for statistical studies of discrimination based on the crude proxies for job qualifications recorded in company records.

ADDITIONAL REFERENCES

- CONWAY, D. A. and ROBERTS, H. V. (1986a). Regression analyses in employment discrimination cases (with discussion). In *Statistics and the Law* (M. H. DeGroot, S. E. Fienberg and J. B. Kadane, eds.) 107-195. Wiley, New York.
- CONWAY, D. A. and ROBERTS, H. V. (1986b). Homogeneous job groups, candidate pools and study of employment discrimination. *ASA Proc. Survey Res. Methods Sec.* Amer. Statist. Assoc.
- CONWAY, D. A. and ROBERTS, H. V. (1987). Analysis of employment discrimination through homogeneous job groups. Working paper, School of Business, Univ. Southern California.
- FIELD, H. F., KIMPTON, M. T., MCGEE, D. W. and ROBERTS, H. V. (1978). Statistical study of equality of employment opportunity at United Airlines: Methods and findings. Report 7824, CMSBE, Graduate School of Business, Univ. Chicago.
- ROBERTS, H. V. (1979). Harris Trust and Savings Bank: An analysis of employee compensation. Report 7946, CMSBE, Graduate School of Business, Univ. Chicago.
- ROBERTS, H. V. (1985a). Direct testimony of defendant's expert Dr. Harry V. Roberts and response of Dr. Roberts to direct testimony of plaintiff's expert Dr. Katherine Monti. *Fauteck v. Montgomery Ward & Co., Inc.*, United States District Court for the Northern District of Illinois, Eastern Division, June 1985.
- ROBERTS, H. V. (1985b). Harris Trust and Savings Bank: An analysis of alleged employment discrimination during the time period 1973-1977; supplement for 1965-1972. Report for Administrative Hearing of *OFCCP v. Harris Bank*, 1985.