

DISCERNING DISCRIMINATION: DOES INTERVIEWING FIRMS MAKE A DIFFERENCE?

Phillip Nelson

State University of New York at Binghamton

and

Solomon W. Polachek

State University of New York at Binghamton

The thread unifying these three symposium papers is that each surveys firms directly to seek information concerning discrimination. Paper one (Riach and Rich) explores the role of gender when firms choose candidates for job interviews; paper two (Moss and Tilly) examines the applicant characteristics (including race) that firms value in the interview process; and finally paper three (Mellor and Paulin) analyzes promotion and wage growth from employee records of two branches of a particular firm in the finance industry. Surveying firms is innovative because most current studies rely solely on worker responses. Because firm surveys are innovative it is not surprising that many of the problems entailed in their use have not been resolved. Some writers allege that direct surveys of firms eliminate biases because they permit crucial productivity measures omitted from traditional employee surveys to be included. Rather than examine the virtues of these, and a whole slew of other papers of the same genre, we concentrate on two questions: Do these studies do better than previous studies to prove the existence of discrimination, as Pressman claims in the introduction? We argue, no. Our point is to show that analyses surveying *only* firms can suffer from the same omitted variable biases inherent in employee surveys. The other question explored: How can similar studies be designed to reveal more about discrimination?

The Mellor-Paulin study makes a quite limited contribution to the literature. Although it surveys firms rather than individuals, it merely applies the methodology used in traditional studies, namely to regress labor market outcome measures on individual characteristics. The innovation is to add a categorical variable representing how one's supervisor ranks each worker's performance. Of course, the problem with this worker performance variable is the same as in related prior studies (e.g., Medoff and Abraham [1980]). Since supervisor ratings can vary by supervisor, what constitutes good work according to one supervisor might constitute poor quality to another. Similarly supervisor rankings can vary across jobs so that the criteria for ranking, say, a key punch operator might vary from those for ranking a statistical analyst. Further, rankings using only five categories, for which literally no one receives rankings below "three," hardly permit the fine differences one would think necessary to evaluate performance fully. Finally, either the authors somehow fail to

obtain or fail to report data on supervisor characteristics, so it is impossible to test whether it makes a difference for a worker to be ranked by a male, a female, a black, a Hispanic, an Asian, a Jew, a Catholic, etc. Without these additional variables the Mellor-Paulin employer-based study provides little that is not already known from surveys of worker wage growth, and it suffers in comparison from much smaller sample sizes of individuals and occupations. Nevertheless, it is interesting that despite, or perhaps because of, these potential difficulties,¹ the authors find *no* net discrimination against one group or another.²

The Moss-Tilly paper is potentially more interesting because it at least uses survey information on how firms *themselves* behave rather than information on worker achievements, which are largely known anyway from traditional employee-based surveys. The study is marred by the choice of an extremely non-random selection of particularly low-skilled Detroit and Los Angeles retail clothing stores, insurance companies, public sector agencies (including utilities and hospitals) and largely black-owned auto parts manufacturers. However, the findings of the paper are threefold. (1) With the exception of automobile parts manufacturers, soft skills such as motivation and ability are more important than "hard" skills such as literacy and being a high school graduate. (2) Blacks are viewed as less likely to have appropriate job skills than whites though it is not clear that firms were asked to compare "comparable" blacks to "comparable" whites, so the reader does not know if these perceptions reflect accurate portrayals or misperceptions. (3) Firms with lower skill requirements and more "negative" racial views have higher proportions of blacks. At least two idiosyncracies mar the interpretation of these results.

First, the labor market described by Moss and Tilly exhibits what search theorists call "search" goods and "experience" goods.³ Search goods most represent jobs requiring hard skills, while experience goods represent jobs requiring soft skills. Firms with hard skill requirements should rely more on testing; whereas firms with soft skill requirements require more informal search procedures such as interviews or referrals. Of the surveyed firms, auto parts manufacturers most rely on hard skills, implying testing. Retail clothes outlets rely on soft skills, implying interviews or referrals (Table 3). Yet contrary to expectations, the Moss-Tilly retail clothes outlets least used employee referrals (Table 7), while auto parts manufacturers used employee referrals most, and written tests least!

One can even use search theory to show that Moss and Tilly understate the importance of hard skills as a barrier to employment. In the real world workers apply for jobs where their probability of obtaining employment is greatest. Since hard skills can be more easily determined by both firms and workers, workers are more likely to self-screen for those attributes. Self-screening comes from two sources: (1) workers have better information about their own hard skills, and (2) workers' chances of success in fooling firms about their hard skills are smaller because firms can easily detect hard skill deficiencies. Neglecting worker self-screening and concentrating only on firm behavior thus can cause one to overestimate the importance of soft-screening over hard-screening. This neglect of worker self-screening might explain one of the anomalies of the Moss-Tilly results: a higher proportion of blacks in the jobs where

soft skills are important in spite of employer complaints that blacks are particularly deficient in soft skills. If the ratio of blacks to whites deficient in hard skills is higher, blacks are less likely than whites to apply for the hard-skill jobs.

Secondly and equally puzzling are the results on race (Table 10, Model 6). Firms that view blacks more negatively hire blacks more frequently, while (in their sample) heavily black-owned auto parts manufacturers hire blacks the least, *ceteris paribus*. Yet sociological "social contact" theories [Allport, 1954] imply a positive correlation between attributes and number of blacks. As the number of blacks rises within an organization, managers recognize their value and negative attitudes diminish. Consistent with this theory, one would think black-owned firms would discriminate the least. Yet it appears that the largely black-owned auto parts manufacturers sampled by Moss and Tilly discriminate the most. As such, this paper finds that blacks discriminate most against blacks!⁴

Perhaps the Riach-Rich paper is the most interesting because it represents a new and possibly growing genre of research, what the authors call "correspondence testing,"⁵ and what the Urban Institute calls "employment audits." Because of the increasing prevalence of this technique in the literature, getting the methodology straight is important. The correspondence-testing/audit-technique is simple. Two individuals, a black and a white or a male and a female, are matched for all personal characteristics other than race or gender. The individual testers then apply for a job, apply for a housing loan, or merely seek to purchase or rent a dwelling or other commodity. Discrimination is defined as a situation in which firms hire minorities with a relatively low frequency. Examples of such studies include Wienk [1991]; Yinger [1986]; Cross et al. [1990]; Turner, Fix, and Struyk [1991]; and Ayres [1991]; these are all described in a survey article [Fix and Struyk, 1991].

Heckman and Siegelman [1991] devote much detail to technical issues surrounding this statistical methodology, such as how the null hypothesis of "no discrimination" is stated, whether the types of jobs applied for are random, and whether the actors serving as audit pairs are really the same in all including unmeasured ways. Rehashing their comments here is unnecessary; one need only refer to the Heckman and Siegelman paper. Indeed, Riach and Rich satisfy many of Heckman and Siegelman's criticisms concerning audit pairs because letters of application and vitae were used instead of actors. Our point is different, and perhaps methodologically more important, because it is conceptual in nature. We argue that current audit type studies are guilty of the same omitted-variable type biases which led to their development in the first place. Audit studies look only at one aspect of the market — in the case of Riach and Rich, employers' decisions to interview workers. Whereas traditional discrimination studies are allegedly supply oriented and thus guilty of neglecting demand, these studies look at one aspect of employer demand and are guilty of neglecting supply. Let us explain by an example.

To get a job (or interview) workers search for a firm, and firms for a worker. Both workers and firms are efficient in that each seek to minimize search costs. Assume all workers to be indistinguishable in all respects, except one clearly discernible characteristic — for example, whether they live in-state. Assume most job applicants are

in-state residents. Further assume that, for any one of many possible reasons, risk-averse workers apply for an inordinate number of jobs to guarantee at least one offer. The firm receives a minority of out-of-state applications. Clearly the probability that an out-of-state applicant will accept a job offer is low, if for no other reason than high fixed moving costs. If the firm finds it costly to offer a job that is refused, prudent employers will tend to offer jobs only to in-state residents, rather than an out-of-stater with otherwise identical characteristics. And this is efficient since it minimizes costs. Even if the firm knew nothing about individual moving costs, it might assume that the probability that an out-of-state applicant will accept an offer is related to the proportions of out-of-state applicants as well as the proportion of out-of-state workers actually in the firm.

The Riach and Rich correspondence test has nothing to do with out-of-staters, but it does have something to do with gender. Like the case of out-of-staters, a firm can easily conceive the probability of an offer being accepted to be related to the proportion of applicants of a particular identifiable group. For example, relatively few female applicants (all else constant) can easily signal low acceptance probabilities and hence high hiring costs associated with keeping jobs vacant too long.

What do these hiring costs have to do with the correspondence studies? The Riach and Rich study (and others of the same genre) omit at least one crucial supply-side variable: the relative proportion of male and female job applicants. Omitting this variable leads the researcher to expect "comparable" male and female applicants to be equally desirable. However, gender differences in job acceptance probabilities are not considered. In short, the demand-based correspondence-testing/employment-audit studies neglect supply-side considerations. While supply-side analyses might be guilty of omitting demand considerations, clearly demand-side analyses are guilty of omitting supply-side considerations. Thus, whereas one might be reluctant to rely on supply models, one should be equally reluctant to rely solely on demand models.

The other problem with the Riach and Rich data is that it is not clear whether or not wages are negotiable in the jobs they examined. If wages are negotiable, the Riach and Rich procedure could show no discrimination where, in fact, there was lots, but this discrimination would show up in wage differences rather than in probability of employment.

Riach and Rich's correspondence test can be extended to answer important questions. For example, one can better discern discrimination by realizing that the more *relevant* is the information conveyed in the application letters, the less is the statistical discrimination. Since interrupted labor-force participation is the key to most observed gender wage-differences [Polachek and Siebert, 1993], one need only vary the potential for interrupted work careers in the job applications, and then compare the results of their correspondence procedure for men and women who have both had a long history of continuous work with the results on men and women who have not. One can make a case for discrimination if firm choices are more dependent on gender than expected participation.

SUMMARY

Supply-based analyses of discrimination using workers as the unit of observation are often criticized for omitting demand considerations. Moving to firms as the unit of observation is clearly an alternative, and thus this symposium's studies should be commended for adopting this perspective. Our point, however, is that studies that rely on firm data may have pitfalls as well, namely omitting supply considerations. Indeed, our comments illustrate such potential biases in each study. These concerns should not be taken lightly because neglecting them can cause one to severely misestimate discrimination.

NOTES

We wish to thank Francine Blau for valuable discussion.

1. We ignore obvious other biases such as changing the specification of Branch B by condensing the categorical gender/race variables from three to one; and interpreting the coefficient of the "lambda" adjusted promotion variable to represent wage growth for those promoted, when in fact it represents wage growth for those not predicted to be promoted.
2. They find no statistically significant race/gender wage growth differences. Tables 2A and 2B indicate no statistically significant race/gender promotion differences, and no statistically significant male wage growth differences between the exempt and non-exempt sectors. The only statistically significant differences are contradictory. Wage growth and promotion are inversely proportional to an occupation's female incumbents in Branch A, while directly proportional to an occupation's female incumbent in Branch B. In short, there is no overt discrimination either in Branch A or B, but Branch A discriminates against women and Branch B discriminates against men in highly female jobs.
3. Nelson [1970] distinguishes between "search" and "experience" goods. For search goods, attaining price information is more important than obtaining quality information, as the quality of a search good (such as a gallon of gas) is relatively uniform. However, for experience goods (such as a restaurant meal or can of tuna) quality is paramount. Since for search goods price information is harder to obtain than quality information while for experience goods quality information is what's hard to get, theory dictates different search procedures for each type of good. One searches extensively in traditional ways by checking dealer prices and looking at advertisements for search goods, while one most efficiently searches using guide books, friends, and time-intensive strategies such as trying out the good, when one purchases experience goods.
4. Model 6 in Table 10 indicates that auto parts manufacturers employ the fewest blacks. Yet according to Table 1 auto parts manufacturers comprise one-third of the sample. One can infer that unless some very large white auto parts manufacturers discriminate much more heavily than the other auto parts manufacturers, black-owned firms must discriminate in employment of blacks, since 11 percent of the entire sample is black-owned and according to the text most auto parts manufacturers were black owned. Moss-Tilly could have made this job of inferring the behavior of black-owned firms unnecessary by explicitly including a black ownership variable.
5. This should be distinguished from "correspondence analysis," used in psychometrics and sociology which is a multivariate technique for converting cross-tabular data into graphical displays called maps. See Greenacre and Blasius [1994].

REFERENCES

- Allport, G. *The Nature of Prejudice*. Cambridge, MA: Addison Wesley, 1954.
- Ayres, I. Fair Driving: Gender and Race Discrimination in Retail Car Negotiations. *Harvard Law Review*, 104(4), 1991, 817-72.
- Cross, H. K., Mell, J. G., and Zimmermann, W. *Employment Hiring Practices: Differential Treatment of Hispanic and Anglo Job Seekers*. Washington, D.C.: The Urban Institute Press, 1990.
- Fix, M. and Struyk, R. An Overview of Auditing for Discrimination. Paper Presented at the Conference on Testing for Discrimination in America. Washington, D.C., September 1991.
- Greenacre, M. and Blasius, J. *Correspondence Analysis in the Social Sciences*. San Diego: Academic Press, 1994.
- Heckman, J. and Siegelman, P. An Economic and Econometric Examination of the Audit Methodology. Paper Presented at the Conference on Testing for Discrimination in America. Washington, D.C., September 1991.
- Nelson, P. Information and Consumer Behavior. *Journal of Political Economy* 78(2), 1970, 311-29.
- Polachek, S. and Siebert, S. *The Economics of Earnings*. Cambridge: Cambridge University Press, 1993.
- Turner, M., Fix, M. and Struyk, R. *Housing Discrimination in America: Summary of the Findings of the Housing Discrimination Study*. Washington, D.C.: The Urban Institute, 1991.
- Weink, R., Reid, C. E., Simonson, J. C. and Eggers, F. J. *Measuring Racial Discrimination in American Housing Markets: The Housing Market Practices Survey*. Washington, D.C.: Office of Policy Development and Research, U.S. Department of Housing and Urban Development, 1991.
- Yinger, J. Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act. *American Economic Review*, December 1986.