statistical discrimination

Abstract:
Statistical discrimination is a theory of inequality between demographic groups based on stereotypes that do not arise from prejudice or racial and gender bias.When rational, information-seeking decision makers use aggregate group characteristics, such as group averages, to evaluate individual personal characteristics, individuals belonging to different groups may be treated differently even if they share identical observable characteristics in every other aspect. Discrimination can be the agents' efficient response to asymmetric beliefs, or discriminatory outcomes may display an element of inefficiency: the disadvantaged group could perform better if beliefs were not asymmetric across groups (but beliefs are asymmetric because the disadvantaged are not performing as well as the dominant group).

Statistical discrimination is a theory of inequality between demographic groups based on stereotypes that do not arise from prejudice or racial and gender bias. It occurs when rational, information-seeking decision makers use aggregate group characteristics to evaluate relevant personal characteristics of the individuals with whom they interact. Because group-level statistics, such as group averages, are used as a proxy for the individual variables, individuals belonging to different groups may be treated differently even if they share identical observable characteristics in every other aspect.
Some examples may help clarify the concept.

* When an insurance company chooses life insurance premia, the customer's likelihood of dying is one of the most relevant variables affecting the company profitability. A seemingly irrelevant proxy, gender, is highly correlated with death frequencies at every age. It is therefore optimal for the company to adopt a policy setting different premia for men and women who share similar characteristics.
* Employers usually place value on job attachment, for reason such as the costs of specific human capital investment. Historically, women have had lower labour market attachment than men, perhaps because of a higher propensity to be involved directly in child-rearing. In evaluating workers with otherwise identical characteristics, employers may prefer hiring male over identical female candidates. This is because employers assess probabilistically higher profitability from hiring a man.
* Highway police are often accused of searching cars driven by minorities more frequently than other cars. While such a policy may be viewed as unfair, it may not be the outcome of prejudice if police officers hold (perhaps biased) beliefs that minorities are more likely to engage in criminal activities. They use such beliefs to maximize the probability of arrest over a given time frame.

All of the above examples share the following features: the decision maker a) is a rational utility maximizing agent engaged in perfecting the available information; b) has incomplete information about some outcome-relevant individual characteristics; and c) holds asymmetric beliefs regarding the average value of relevant variables across groups. These beliefs can be interpreted as stereotypes.

The examples are also different in one important aspect. In the first example, differences in the company's beliefs are rooted in a 'technological', exogenous difference between groups, the death frequencies. In the other examples, asymmetric beliefs may feed back into differences in the type of behaviour that generates them. To clarify: it is possible that, because employers are less likely to hire women in jobs that require labour market attachment, then women are more likely to be involved in child-rearing than men, and are less prone to acquire the skills that are necessary to seek and perform well in those jobs, confirming the asymmetric belief that employers hold regarding labour market attachment. In this case, beliefs are endogenous, or self-confirming.

The distinction between these two sources of inequality is important. In the first case, discrimination is the agents' efficient response to asymmetric beliefs. In the second case, discriminatory outcomes may display an element of inefficiency: the disadvantaged group could perform better if beliefs were not asymmetric across groups (but beliefs are asymmetric because the disadvantaged are not performing as well as the dominant group).

We proceed by presenting in greater detail two additional labour-market-related examples of these two flavours of statistical discrimination.

Discrimination when the quality of information differs exogenously across groups

Consider the example of an employer that does not observe with certainty the skill level of her prospective employees. The population of workers of a given group has a skill distribution Φ. Workers draw from Φ their skill p, which is assumed to be equal to the value of their product when employed. Employers know the distribution, but only observe a noisy signal of productivity, $s = p + e$, where e is a zero-mean error distributed according to Φ_e. The employer infers p from s using the available information. In equilibrium, if the labour market is competitive and all employers share the same type of information, workers are paid according to their expected productivity conditional on the value of the signal. It can be shown that if Φ is a normal distribution, with mean μ and standard deviation σ, and if the error is also normal, with mean 0 and variance σ_e, then, using Bayes's rule, the employers' best estimate of p is a weighted average of the workers' signal and the population average, with weights that depend on the relative size of the variance of the skill and the error distribution (see Phelps, 1972). Formally, the expected worker's productivity conditional on the signal is:

$$E(p \mid s) = \frac{\sigma_e^2}{\sigma^2 + \sigma_e^2} p + \frac{\sigma^2}{\sigma^2 + \sigma_e^2} s \tag{1}$$

Intuitively, if the signal is very noisy (that is, if the variance of e is very high), the expected conditional value of workers productivity is close to the population average regardless of the signal's value. On the other extreme, if the signal is very precise (σ_e is close to zero), then the signal provides a precise estimate of the worker's ability.

If workers belong to two identifiable demographic groups, then it is rational for the employer to condition her inference also on the group identity of the worker. Suppose for example that the signal emitted by minorities is 'noisier' than the signal of non-minorities (perhaps because tests are race-biased). Then it follows that minorities with high signals will receive lower wages than same-signal workers from the dominant group, and the opposite happens to workers with low signals.

While this model is capable of explaining differential treatment for same-signal workers from different groups, on average workers of the two groups receive the same wage, which is equal to the average productivity μ. Group differences between groups' average wages can be obtained by extending the model to include employers' risk aversion (Aigner and Cain, 1977), or workers' pre-market investment in human capital (Lundberg and Startz, 1983). All of these approaches still require the assumption of some form of exogenous group difference, for instance in the signal's quality, an assumption which has been questioned.

Equilibrium discrimination with *ex-ante* identical groups

Recognizing these limitations, Arrow (1973) proposed an alternative model where inequality occurs even with identical groups' fundamentals. In his model, employers' asymmetric beliefs about members of different groups are self -confirming.

A formalization of this approach is presented in Coate and Loury (1993). There are two job-tasks, a simple task that anybody can perform, and a complex task requiring skills that can only be acquired through prior costly investment in human capital. Workers are heterogeneous in the cost of investment, but the cost distribution is the same across groups. There is a linear technology: each worker has the same productivity in the low-skill job, but workers have zero productivity in the high-skill job if they made no investment in human capital. If they do invest, their productivity is higher in the high-skill job than in the low-skill job. Wages are set exogenously and the high-skill job is paid with higher wages than the simple job.

Employers do not observe skill level, and assign workers to tasks according to an imperfect signal that is correlated with investment (this may be viewed as the outcome of a test where the likelihood of receiving a higher grade is higher for workers that have invested in human capital). No differences in the quality or informativeness of the test across groups need to be assumed. The task assignment depends on the expected skill, which depends both on the individual signal and on the group average, for reasons identical to those outlined in Phelps's model above, and formalized in equation (1). The optimal task assignment rule is to set a threshold and assign all workers with signals above such a threshold to the high-skill job, and workers below that threshold to the low-skill job. The marginal worker, with a signal equal to the threshold, has an expected productivity in the high-skill job identical to her productivity in the low-skill job.

A crucial feature of the model is that human capital investment provides a positive informational externality: investing not only helps a worker's own chances of being assigned to the high-skill job, but also increases the probability of being qualified to perform the high-skill job for every member of her group. This externality may generate multiple equilibria with different fractions of aggregate human capital investment.

For an intuition on how equilibria may be characterized, consider a model with only one group of workers. Suppose that in equilibrium only a few workers acquire human capital. In this case, employers' assessment of the probability that a worker has invested in human capital, given her signal, will be low even if the signal is relatively high. This is because Bayes's rule implies that the posterior probability of one worker having invested in human capital is increasing with the prior – the group's aggregate investment, (in equation (1) the workers' expected productivity depends on the group's average). Therefore, the optimal threshold for task assignment is set relatively high. Because it is difficult to obtain a high-skill job, the expected wage gain from investing is small, so only the few workers with relatively low cost of investment will acquire human capital, which confirms the original assumption that few workers invest.

There may also be equilibria with a high fraction of workers acquiring human capital. In such equilibria, employers will set a lower threshold in order to assign workers to a high-skill job relative to the low-investment equilibrium. Because high-skill jobs are more accessible, this task assignment rule provides higher returns to investment in human capital, which is a necessary condition to support this as an equilibrium. Note that the threshold cannot be too low, because if the high-skill job becomes too easily accessible, there are no incentives to invest in human capital.

Turning to the two-group model, all outcomes where both groups replicate one of the equilibria of the one-group model are also an equilibrium of the two-group model. These symmetric equilibria display no inequality. However, if the base model displays multiple equilibria, and if groups coordinate on equilibria with different fractions of workers investing in human capital, the group with lower investment will exhibit lower average wages, a higher fraction of workers employed in the low-skill job, and a higher threshold required for assignment to the high-skilled job.

While such a model illustrates the possibility for inequality to arise even when groups are *ex ante* identical, it cannot predict which group will be discriminated against, or why a symmetric equilibrium was not selected. The linearity of the technology implies that groups are treated separately, as if they were living in different islands: expected marginal productivities of workers depend only on their own signal and on the aggregate investment of their own group. Therefore, in this environment, statistical discrimination exists because of a *coordination failure*: the disadvantaged group fails to coordinate on the 'good' equilibrium, but the dominant group has nothing to lose if the disadvantaged group could solve the coordination failure.

A version of the model with a more general technology provides an alternative source of discrimination in which groups have conflicting interests (Moro and Norman, 2004). Consider a production function exhibiting a complementarity between tasks. Then, the marginal product of a worker in each task is affected by aggregate investment in human capital in both groups. Specifically, the expected marginal product in the high-skill job of a given worker depends negatively on aggregate investment in human capital from members of the other group. This is because when more members of the other group acquire human capital, the higher aggregate availability of skills decreases the marginal product of a skilled worker. Hence, incentives to acquire skills decrease when more members of the other group acquire skills. The complementarity generates incentives for groups to *specialize,* and asymmetric equilibria may exist even if there is a unique symmetric equilibrium. While there is an element of self-fulfilling prophecy, asymmetric equilibria here are the result of specialization rather than coordination failure. In such equilibria, the discriminated group cannot coordinate on a better outcome without a simultaneous coordination on a worse outcome by the other group: the dominant group always gains from discrimination. While in this model there always exist symmetric equilibria, group size is a relevant factor, and the roles of the two groups can be reversed only if group sizes are identical.

Empirical evidence of statistical discrimination

There is a vast body of literature documenting racial and gender wage inequality. (See the bibliographies for the articles in this *Dictionary* on black-white labour market inequality in the United States, and women's work and wages.) In such literature, group differences that cannot be explained by differences in observable characteristics are attributed to prejudicial preferences, and little attention has been devoted to testing whether statistical discrimination plays a role in determining such differences. The main problem is to find ways to identify, using available data, to what extent group differences are caused by prejudicial attitudes, or by asymmetric beliefs (self-confirming or otherwise) and incentives.

Altonji and Pierret (2001) observe that if firms statistically discriminate, then as firms learn over time about workers' productivity, differences on the observed variables should fall over time. The data supports this proposition. Another example of an 'outcome-based' approach in the identification of the source of discrimination is the study by Knowles, Persico and Todd (2001), testing discrimination against minorities in motor vehicle searches by police officers. They find no evidence of racial animus in the data.

Attempts to identify different sources of discrimination include experimental or quasi-experimental data. (See Anderson, Fryer and Holt, 2006 for a survey of the literature, which also includes sources from the psychology literature.)

A different approach is to estimate statistical discrimination models directly. Moro (2003) finds that adverse equilibrium selection did not play a role in exacerbating wage inequality during the last part of the 20th century. Fang (2006) estimates a statistical discrimination model to assess the prevalence of a signalling component to the college wage premium. While the estimates match wage distributions reasonably well, they are not designed to answer questions about model validation.

Andrea Moro

See also **affirmative action; black–white labour market inequality in the United States; racial profiling; taste-based discrimination; women's work and wages.**

Bibliography

Aigner, D. J. and Cain, G. 1977. Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review* 30, 749–76.

Altonji, J. G. and Pierret C. R. 2001. Employer Learning And Statistical Discrimination. *The Quarterly Journal of Economics* 116(1), 313-350

Arrow, K. J. 1973.The theory of discrimination. In *Discrimination in Labor Markets*, ed. O. Ashenfelter and A. Rees. Princeton, N.J., Princeton University Press.

Anderson, L., Fryer, R. and Holt, C. 2006. Discrimination: experimental evidence from psychology and economics. In *Handbook on Economics of Discrimination*, ed. W. Rogers. Cheltenham, UK/Northampton, Mass.: Edward Elgar.

Coate, S. and Loury, G.C. 1993. Will affirmative action policies eliminate negative stereotypes? *American Economic Review* 83, 1220–40.

Fang, H. 2006. Disentangling the college wage premium: estimating a model with endogenous education choices. *International Economic Review* 47(4), 1151–85.

Knowles, J., Persico, N. and Todd, P. 2001. Racial bias in motor vehicle searches: theory and evidence. *Journal of Political Economy* 109(1), 203-29

Lundberg, S. J. and Startz, R. 1983. Private discrimination and social intervention in competitive labor markets. *American Economic. Review* 73, 340–7.

Moro, A. 2003. The effect of statistical discrimination on black-white wage inequality: estimating a model with multiple equilibria. *International Economic Review* 44(2), 457–500.

Moro, A. and Norman, P. A. 2004. General equilibrium model of statistical discrimination. *Journal of Economic Theory* 114(1), 1–30.

Phelps, Edmund S. (1972). The Statistical theory of Racism and Sexism. *American Economic Review* 62(4), 659-61.

Keywords: discrimination, inequality, statistical discrimination, wage determination.