Self-fulfilling Regression and Statistical Discrimination

Takashi Hayashi*

May 11, 2017

Abstract

This paper provides a simple model which explains that statistical discrimination can arise in a purely self-fulfilling manner. The story is: (i) at the point of hiring employers cannot observe workers' productivities but can observe only their signals, such as test score, and under perfect competition they pay expected labor productivity conditional on signal observation, based on their belief about *return to signal*; (ii) given the employers' belief, workers choose effort level, which affects joint probability distribution over pairs of productivity and signal; (iii) in equilibrium, the employers' belief proves to be *statistically consistent*. We show that there may be multiple equilibria, and equilibrium selection has nothing to do with economic fundamental.

*Adam Smith Business School, University of Glasgow. email: `Takashi.Hayashi@glasgow.ac.uk`

1 Introduction

This paper provides a simple model which explains that statistical discrimination can arise in a purely self-fulfilling manner, that is, as a multiple-equilibria phenomenon.

The story is: (i) at the point of hiring employers cannot observe workers' productivities but can observe only their signals, such as test score, and under perfect competition they pay expected labor productivity conditional on signal observation, based on their belief about *return to signal*; (ii) workers choose effort level, which affects joint probability distribution over pairs of productivity and signal, and such effort is high/low when the employers' belief about return to signal is high/low; (iii) in equilibrium, the employers' belief proves to be *statistically consistent*, as one can observe statistical relationship between signal and productivity ex-post. We show that there may be multiple equilibria, and different groups of workers may fall in different equilibria even when they have totally the same fundamental.

Statistical discrimination as a self-fulfilling prophecy phenomenon has been widely studied. The idea of self-fulfilling prophecy in the context of discrimination dates back to Merton [7] and Myrdal [9]. Among the works in economic theory, the prominent ones are Arrow [2] and Coate and Loury [4]. The basic story throughout is that as employers have a belief that a given group has high/low productivity, it encourages/discourages the workers in that group to spend effort on skill formation in some way, and its consequence verifies the employers' belief in a self-fulfilling manner, while the precise driving-force of this cycle differs across the works.

Arrow [2] considers that there are two kinds of jobs, skilled and unskilled, and in order to be qualified for the skilled job workers need to pay some cost, which is drawn from a distribution that is identical across groups. The employers may have a belief that different groups have different proportions of being qualified, and pay different wages for the same skilled job in order offset the cost of verifying skills. Because of the difference of wages, different groups of workers result in generating different proportions of workers whose cost for qualification is cheap enough, which verifies the employers' belief in a self-fulfilling manner. Coate and Loury [4] also consider that there are two kinds of jobs, skilled and unskilled, where the wage of each job is exogenously given. Workers need to pay some cost to acquire the skill, which is drawn from a distribution that is identical across groups. However, the workers' investments are unobservable and the employers cannot verify if a given worker is indeed qualified at the point of hiring. Instead, the employers assign workers to the jobs based on signals, where its distribution conditional on being actually skilled/unskilled is common across the groups. The employers may have a belief that different groups have different proportions of qualified workers, and set different bars on signals for the skilled job. Because of the difference in the bards workers in different groups face different expected wages for the skilled job. Hence different groups of workers result in different proportions of workers who in fact acquire the skill, which verifies the employers' belief in a self-fulfilling manner.[1]

Although much of the ideas are there in those studies, we would emphasize significances of our model as follows. First, what affects workers' effort is employers' belief about *return to signal*, that is, their belief about *slope*, rather than their belief about intercept or marginal belief about regressor (signal) or about regressand (productivity). This contrasts to Arrow [2] and Coate and Loury [4] in which self-fulfilling belief is about proportion of qualified workers. Note that this is not just a technical difference, since belief about intercept alone does not affect workers' investment, unless extra frictions are invited and combined with it, as it determines only the constant

[1] For an extensive survey on statistical discrimination in general, see Fang and Moro [5] and references therein.

term in the formula of expected return of effort.

Second, given employers' belief about return to signal, workers rationally choose effort level which affects joint distribution over productivity and signal. We allow that informativeness of signal is endogenously determined, and show that different degrees of informativeness arise across different equilibria, despite that they are coming from the same underlying structure. This contrasts to the models of regression-based statistical discrimination as presented by Phelps [10], Aigner and Cain [1], in which such effort choice is an implicit fixed factor, and also to Lundberg and Startz [6] which assumes exogenously different degrees of informativeness of signal across groups.

Third, the employers are "rational" and the labor market is "constrained efficient," in the sense that they are paying expected productivity conditional on observation of signals, which is based on their conditional belief being determined in equilibrium, and their belief is *statistically verified* in equilibrium. This contrasts to Arrow [2] in which discriminatory wages have to differ from marginal productivity of labor despite that skills are assumed to be observable, and to Coate and Loury [4] in which wages are assumed to be exogenous.[2]

The paper proceeds as follows. In Section 2 we present the model and the definition of equilibrium. In Section 3, we focus on the case of linear regression, and show that there can be multiple equilibria in which the employers' belief about return to signal is self-fulfilling. In Section 4 we conclude by discussing policy implications.

[2]Moro and Norman [8] extend the model by Coate and Loury so that wages are endogenously determined, maintaining that self-fulfilling belief is about proportion of qualified workers.

2 The Model and Equilibrium Definition

Consider a group populated with a continuum of identical workers. Each worker can spend effort on training or schooling, denoted by $z \geq 0$. A worker's effort affects probability distribution over pairs of actual productivity y and signal x, where the productivity is measured in an output good. Assume that there is no income effect on the output good, and the workers are risk-neutral.

Let $f(x,y)[z]$ be the density of signal/productivity pair (x,y), $f(x)[z]$ be the marginal density of signal x, and $f(y|x)[z]$ be the density of productivity y conditional on signal x, when the effort level is z.

Consider that there are identical employers. Employers' belief is given in the form of density $\phi(x,y)$ which is to be determined in equilibrium. Let $\phi(y|x)$ denote the density of y conditional x, and let $\phi(x)$ denote the marginal density of x. We consider that, ex-post, once can observe the distribution of pairs of signal and actual productivity.

Assume that the labor market is competitive. However, note that the employers cannot observe the workers' productivities at the point of contracting, which can be observed only ex-post. Also, the employers cannot monitor or verify the workers' training efforts. Hence under perfect competition the employers simply pay a given worker expected productivity conditional on signal observation. Thus, a worker receives

$$\int y\phi(y|x)dy$$

when she sends signal x.

Let $C(z)$ be the cost of effort z. Then a generic worker solves

$$\max_z \int \int y\phi(y|x)dyf(x)[z]dx - C(z).$$

Denote the solution by $z(\phi)$, as the workers' optimal choice depends on the employers' belief ϕ.

In equilibrium the employers' belief is required to be *statistically consistent*, that is,

$$\phi(x,y) = f(x,y)[z(\phi)]$$

must hold. This equilibrium condition may be written in two equations,

$$\begin{aligned}\phi(y|x) &= f(y|x)[z(\phi)] \\ \phi(x) &= f(x)[z(\phi)],\end{aligned}$$

and the first condition is what matters and the second is rather a consequence of it.

In the next section we present a simple example of multiple equilibria and explain the source of multiplicity. The same point can be illustrated instead with a model in which productivity level is binary and signal level is binary, for example. We chose the linear-normal model there in order to relate to the linear regression practice as done in applications broadly.

3 Self-fulfilling Linear Regression

Consider that joint density of signal and productivity given effort level z follows normal distribution,

$$(x,y) \sim N\left((\mu_x(z),\mu_y(z)), \begin{pmatrix} \sigma_x^2(z) & \rho(z)\sigma_x(z)\sigma_y(z) \\ \rho(z)\sigma_x(z)\sigma_y(z) & \sigma_y^2(z) \end{pmatrix}\right).$$

It yields marginal density of signal

$$x \sim N(\mu_x(z), \sigma_x^2(z))$$

and conditional density

$$y|x \sim N\left(\mu_y(z) - \frac{\sigma_y(z)}{\sigma_x(z)}\rho(z)\mu_x(z) + \frac{\sigma_y(z)}{\sigma_x(z)}\rho(z)x, (1-\rho^2(z))\sigma_y^2(z)\right).$$

Then the natural candidate of equilibrium belief is also a normal distribution. Let us take the regression form to describe this belief, that is,

$$y = \alpha + \beta x + \varepsilon, \quad \varepsilon \sim N(0, \gamma^2)$$

and

$$x \sim N(\nu_x, \gamma_x^2).$$

Then, since

$$\int y\phi(y|x)dy = \alpha + \beta x$$

and

$$\int (\alpha + \beta x) f(x)[z] = \alpha + \beta\mu_x(z),$$

a generic worker solves

$$\max_z \alpha + \beta\mu_x(z) - C(z).$$

Assume monotonicity and concavity condition

$$\mu_x'(z) > 0, \quad \mu_x''(z) \leq 0$$

and

$$C'(z) > 0, \quad C''(z) > 0$$

for all $z > 0$, and boundary condition

$$\lim_{z\to 0} C'(z) = 0, \quad \lim_{z\to\infty} C'(z) = \infty,$$

or one can allow constant marginal cost by assuming the corresponding strict concavity and boundary condition on μ_x.

Then the optimal choice is given by the first-order condition

$$\beta\mu_x'(z) = C'(z).$$

Denote its solution by $z(\beta)$, as it depends only on β and not on α or γ or else, and from the implicit function theorem it holds

$$z'(\beta) = \frac{\mu_x'(z(\beta))}{C''(z(\beta)) - \beta\mu_x''(z(\beta))} > 0.$$

Now the statistical consistency condition in equilibrium imposes

$$\begin{aligned}
\alpha &= \mu_y(z(\beta)) - \frac{\sigma_y(z(\beta))}{\sigma_x(z(\beta))}\rho(z(\beta))\mu_x(z(\beta)) \\
\beta &= \frac{\sigma_y(z(\beta))}{\sigma_x(z(\beta))}\rho(z(\beta)) \\
\gamma^2 &= (1-\rho^2(z(\beta)))\sigma_y^2(z(\beta)) \\
\nu_x &= \mu_x(z(\beta)) \\
\gamma_x^2 &= \sigma_x^2(z(\beta)).
\end{aligned}$$

Since there is no circularity in the equations except the one for β, once β is determined all of $\alpha, \gamma, \nu_x, \gamma_x$ are determined uniquely.

Now let us focus on the fixed-point condition

$$\beta = \frac{\sigma_y(z(\beta))}{\sigma_x(z(\beta))}\rho(z(\beta)) \equiv H(\beta).$$

Notice that no restriction on $\sigma_y(\cdot)$, $\sigma_x(\cdot)$ and $\rho(\cdot)$ is needed to ensure the workers' optimal choice problem to have a smooth interior solution $z(\cdot)$, as it depends only on $\pi_x(\cdot)$, $C(\cdot)$. Moreover, no restriction on higher-order derivatives of $\pi_x(\cdot)$, $C(\cdot)$ other than on their first and second derivatives is made, hence there is no restriction on $z(\cdot)$ besides it is increasing. Hence it is easy to have H to exhibit a graph like in Figure 1. In the figure there are two equilibria, $\underline{\beta}$ and $\overline{\beta}$, which are stable with respect to adaptive learning. However, which equilibrium is selected has nothing to do with economic fundamental. Therefore, one group of workers may fall into $\underline{\beta}$ and another group of worker may fall into $\overline{\beta}$, even if they have totally the same fundamental. As $z(\cdot)$ is

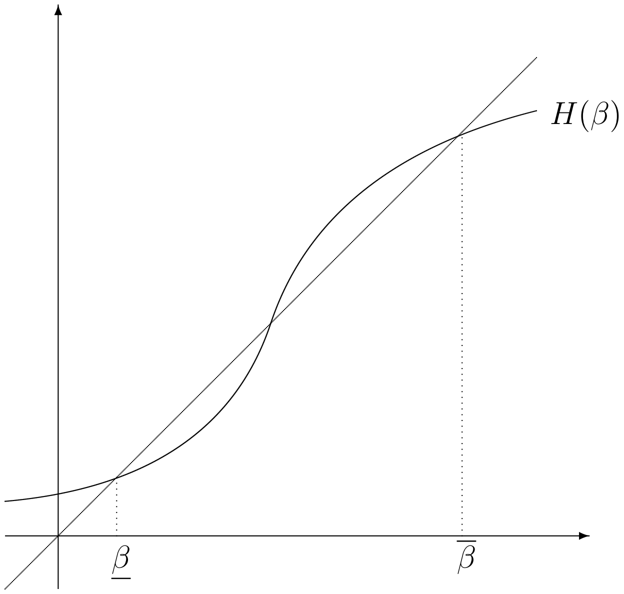

Figure 1: Fixed-point mapping for β

increasing and $\pi_x(\cdot)$ is increasing, $\underline{\beta}$ results in lower mean of observed signal and $\overline{\beta}$ leads to higher mean of observed signal. Therefore, a group of workers in a low equilibrium yields low return to signal and exhibits low distribution of signal as well, and a group of workers in a high equilibrium yields high return to signal and exhibits high distribution of signal as well.

Since the employers have zero expected profit in all equilibria, because of perfect competition, all of expected net surplus

$$V(z) = \mu_y(z) - C(z)$$

goes to the worker side. Hence all equilibria are Pareto-ranked. As far as the effort levels in equilibria are below the first-best level, higher equilibrium generates higher welfare and lower equilibrium generates lower welfare, although it is possible that return to signal is too much overrated and leads to inefficiency in the opposite side.

4 Conclusion

Of course we do not claim that the self-fulfilling nature as demonstrated above is the only cause of statistical discrimination or discrimination in general. However, it gives us the following lessons.

First, if our point indeed applies to the actual discriminations, unless employers' belief is drastically altered just promoting opportunities of training will not help. And, even if a policy can affect employers' belief to change, such change may not result in a better one or the intended one. This is consistent with the discussion by Coate and Loury [4].

Second, because the employers are in fact profit-maximizing under the given kind of market incompleteness and informational asymmetry, they will not be driven out by means of market force at least easily, even if they have misperception of the underling mechanism of skill formation. In the world written only in contractible terms, being "superficially correct" is enough for profit maximization, and the existing employers cannot be outperformed by potential entrants, since any contingencies about which the latter may have better understandings than the former are not contractible.[3]

References

[1] Aigner, Dennis J., and Glen G. Cain. "Statistical theories of discrimination in labor markets." Industrial and Labor relations review (1977): 175-187.

[3] One may wonder if the multiple-equilibrium phenomenon as above is robust when learning comes in, which is not necessarily adaptive and may be more sophisticated. We might follow the approach by Blume [3], which is adopts Coate and Loury [4] as the static benchmark, considers anticipatory learning and show that the full-employment equilibrium is stochastically stable and the other equilibria are stable in short-run but not in long-run.

[2] Arrow, Kenneth Joseph. "The Theory of Discrimination", in O. Ashenfelter and A. Rees (eds.) Discrimination in Labor Markets (Princeton, NJ: Princeton University Press), 3-33.

[3] Blume, Lawrence E. "Learning and statistical discrimination." The American economic review 95.2 (2005): 118-121.

[4] Coate, Stephen, and Glenn C. Loury. "Will affirmative-action policies eliminate negative stereotypes?." The American Economic Review (1993): 1220-1240.

[5] Fang, Hanming, and Andrea Moro. Theories of statistical discrimination and affirmative action: A survey. No. w15860. National Bureau of Economic Research, 2010

[6] Lundberg, Shelly J., and Richard Startz. "Private discrimination and social intervention in competitive labor market." The American Economic Review 73.3 (1983): 340-347.

[7] Merton, Robert K. "The self-fulfilling prophecy." The Antioch Review 8.2 (1948): 193-210.

[8] Moro, Andrea, and Peter Norman. "Affirmative action in a competitive economy." Journal of Public Economics 87.3 (2003): 567-594.

[9] Myrdal, Gunnar. An American dilemma, Volume 2: The Negro problem and modern democracy. Vol. 2. Transaction Publishers, 1944.

[10] Phelps, Edmund S. "The statistical theory of racism and sexism." The american economic review 62.4 (1972): 659-661.