Stop blaming the victim : credit discrimination without moral hazard

Dan SASAKI

Abstract

Can a specific class of borrowers be treated discriminatorily even if (1) the class is identical with the remainder of the population in terms of their ability and their use of borrowed money, and also (2) lenders are rational and competitive? Unlike most of the existing *statistical discrimination* literature, the endogenous discrimination in this paper does not involve hidden actions which are distinct between the discriminated-against and the discriminated-for. Thereby it is successfully shown that a discriminatory self-fulfilling prophecy can be endogenously sustained even when the class of borrowers do nothing "blameworthy" to contribute to such a prophecy.

Keywords

statistical discrimination, hidden action, moral hazard.

JEL classifications: K13, D61, L11.

Acknowledgements: The author should like to thank Todd R. Kaplan, David DeMeza, Timothy Van Zandt, and the seminar participants at Cardiff Business School INSEAD, and at Econometric Society Australasian Meeting (Auckland, New Zealand, 2001) for discussions and comments on this paper. The author assumes sole responsibility for any remaining imperfections and welcomes further comments.

1. Introduction

STATISTICAL DISCRIMINATION literature, represented by Arrow (1973) and Coate and Loury (1993) *inter alia*, has it that if an *externally recognisable* class (e.g., a race, or a gender) of workers were to expect discriminatorily lower wages than other classes, then such a class would have accordingly lower incentives to invest in their own productive human capital, which in turn would legitimise the wage discrimination as an accurate *self fulfilling prophecy* for the actual productivity of different classes of workers. Thereby in the presence of *multiple equilibria*, there can be a *separating equilibrium* where different classes of workers reside in different equilibria, even if workers from different classes are *a priori identical* in the sense that their externally recognisable features have no inherent effect on their economic productivity. This theory provides a neat, incisive response to its civil-rights-era predecessor, the old assertion that race and gender discrimination could uphold only if employers were economically irrational.

Analogies can be applied to other markets than labour markets as well. In a credit market, as is well known, unfavourable lending conditions can encourage *moral hazard* on the borrowers' side. Thereby if a class of borrowers are treated less favourably than others, then they are likely to use the fund in "hazardous" ways, which in turn justifies the unfavourable treatment from the lenders' side.

According to this theory, it remains technically unclear who is to "blame" in the first place. Namely, since the low human capital investment made by the discriminated-against class and the discriminatory treatment they receive in the economy are complementing each other, both the discriminator and the discriminated-against are contributing to the discriminatory self-fulfilling expectations. Viewed more practically, this leads to a pragmatic question whether discrimination would disappear once all the hidden actions taken by the discriminated-against group were somehow monitored, controlled and corrected appropriately. In other words, can discrimination self-sustain without "cooperation" of the discriminated-against?

This paper probes the possibility that lenders' discriminatory actions alone, without borrowers' moral hazard or any hidden action at all, self-sustain as a competitive equilibrium. The basic model is laid out in section 2. Credit discrimination by denial to lend money is discussed in section 3. Discrimination in terms of interest rates is then discussed in section 4. Section 5 informally probes the seemingly paradoxical prospect that a less than perfectly competitive credit market might eliminate some aspects of discrimination. Possible (counter)effects of legally prohibiting discrimination are discussed in section 6. A concluding remark is provided in section 7.

2. The model

There are two *externally recognisable* classes of borrowers. These two classes are labelled W and B. To run his/her business, a borrower needs to be always borrowing one unit of fund without interruption. The lending term of a loan is nevertheless finite[1)], hence at the maturity of each term the borrower must find a new loan, or else go out of business.

The key observation that motivates this model is the possibility that whether a borrower can repay the previous loan that is about to mature is not necessarily independent of the continuation or discontinuation of his/her business. Theoretically, this influence can materialise in either direction, but perhaps the most interesting case to observe is that the continuation of his/her business facilitates the borrower to repay his/her previous loan.

This translates into the following. At the end of each lending term, a borrower is in one of the three states.

Success (probability s): The loan can be repayed unconditionally, whether the borrower continues his/her business or not.

Refinancible failure (probability f): The loan can be repayed if and only if the borrower continues his/her business.

Insolvency (probability $1 - s - f$): The loan cannot be repayed.

1) It is outside the scope of this paper to question why a loan has a finite term. Nonetheless, the implication of a longer-term relationship maintained by a specific lender-borrower pair shall be very briefly discussed in section 5.

These probabilities *vary across borrowers*. Lenders can, based upon statistically published data, assess each borrower's idiosyncratic probability parameters s and f, based upon the borrower's *objectively verifiable* personal profiles (such as education and criminal records). These parameters are *unaffected*, however, by whether the borrower belongs to class W or class B.

Obviously, given that the continuation of his/her business can enhance a borrower's ability to repay his/her previous loan, an intertemporally efficient scheme would be for the same lender to help the borrower refinance his/her business a term after another, so as to maximise the possibility of its continuation. In reality, however, there is no guarantee that the lending decision this term and the lending decision next term are undertaken by the same individual. What if the lender discontinues lending funds to the specific kind of industry or business? What if, even within the same lending institution, a new fund manager takes over the old one? And what if a fund manager's job hinges upon how many of his/her loans have been *promptly* repayed? The gist here is that, in many of these very realistic scenarios, the lending decisions are not always made in the most intertemporally efficient way.

Fund managers' "myopia" can be taken into account as follows. Each fund manager knows that there is a positive probability that the same borrower will be dealt with by another fund manager. Thereby the probability that the loan can be repayed at its maturity is $s + \nu f$ where $\nu = 1$ if and only if *all* other fund managers are willing to lend a fund to the borrower, and in general ν is strictly increasing in the fraction of other fund managers who are willing to lend to the same borrower.

All fund managers are assumed to be *a priori* identical and, aside from the aforementioned "myopia", competitive. The risk-free interest rate in the economy is denoted by $\underline{r}$ and the interest rate for risky loans is denoted generally by r where $r > \underline{r}$. Throughout the paper, it is assumed for computational simplicity[2)] that no partial repayment can be made. That is, the lender loses the whole amount of the loan if the loan is not fully repayed.

2) It is obvious that this assumption is not at all crucial to any of the *qualitative* findings in this paper.

3. Discriminatory lending denial

Suppose first that fund managers offer an exogenously fixed rate r to all borrowers, so that the only decision for fund managers is to select which borrowers to lend and which borrowers to decline.

Generally, a fund manager should issue a loan if and only if the borrower satisfies

$$(s + \nu f)(1 + r) \geq 1 + \underline{r} \tag{3.1}$$

where ν is, as aforementioned, the probability that the borrower can continue his/her business, which is an increasing function of the fraction of other fund managers who are also willing to lend to this borrower, denoted henceforth by θ, and $\nu = 1$ if and only if the fraction of such fund managers is 1.

In particular, if all other fund managers in the economy apply the identical decision rule A, that is, accepting borrowers if and only if they satisfy $\{s, f\} \in A$ where $A \subset \mathbb{R}^2_+$, then the fund manager should accept the borrower if and only if :

$(s + f)(1 + r) \geq 1 + \underline{r}$ when $(s, f) \in A$, or

$(s + \nu_0 f)(1 + r) \geq 1 + \underline{r}$ when $(s, f) \notin A$

where ν_0 is the probability ν when no other fund manager is willing to lend to the borrower.

Hence the following equilibrium characterisation is obtained.

Proposition 1 : Given a fixed interest rate $r > \underline{r}$, all fund managers' adopting the decision rule A such that

$$(s, f) \in A \quad \forall (s, f) \text{ such that } (s + \nu_0 f)(1 + r) \geq 1 + \underline{r},$$
$$(s, f) \notin A \quad \forall (s, f) \text{ such that } (s + f)(1 + r) < 1 + \underline{r},$$

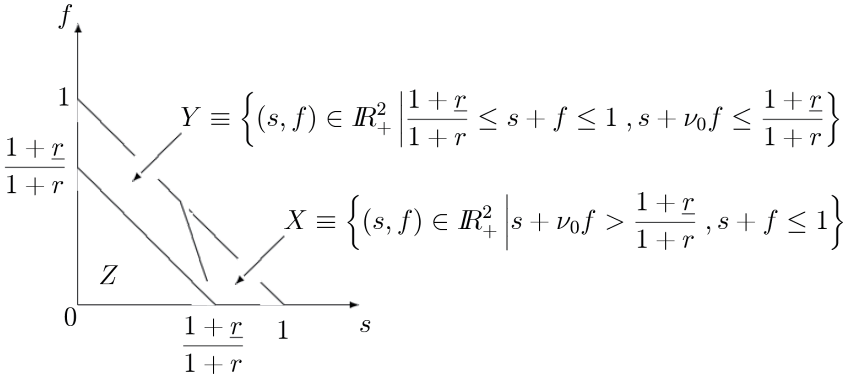
is a competitive equilibrium.

Note that there are many A satisfying these criteria. Figure 1 illustrates the vast multiplicity of equilibria in this model. Any acceptance set A satisfying

$$A \subseteq (X + Y) \quad \text{and} \quad A \supseteq X$$

is a competitive equilibrium.

Figure1: the “must-accept” set (X) and the “may-accept” set (Y).



Intuitively, the region X in figure 1 is the set of those borrowers whose success probability is so high that every fund manager must, as long as economically rational, accept them. The region Z, on the other hand, is the set of those borrowers whose $s+f$ are unacceptably low. The region Y, however, is up to fund managers' discretion in the sense that, if all other fund managers accept a subset of Y then it is the *unique best response* by each fund manager to accept the exact same subset of Y.

Note further that fund managers may apply different acceptance rules for W-class borrowers and B-class borrowers. The following Corollary can be viewed as a generalisation of Proposition 1.

Corollary1: All fund managers' adopting the identical pair of decision rules (A_W, A_B), that is to apply the acceptance set A_W for W-class borrowers whilst applying A_B for B-class borrowers, is an equilibrium if and only if

$$(s,f) \in A_W, \quad (s,f) \in A_B \quad \forall (s,f) \text{ such that } (s+\nu_0 f)(1+r) \ge 1+\underline{r},$$
$$(s,f) \notin A_W, \quad (s,f) \notin A_B \quad \forall (s,f) \text{ such that } (s+f)(1+r) < 1+\underline{r}.$$

In words, the fund managers' discretion over Y can be exercised conditional upon the borrowers' classes. Namely, if all other fund managers impose a discriminatory lending rule by accepting two different subsets of Y for W-class and B-class borrowers, then the *unique best responses* by each fund manager is to adopt the exact same pair of subsets.

Using the same notation as in Figure 1, for example, if all borrowers adopt $A_W = X + Y$ and $A_B = X$, then it is a self-sustained competitive equilibrium that discriminates for class W and against class B.

4. Discriminatory lending rates

Now consider endogenising the interest rate r. The rationality condition for a fund manager's decision is basically the same as (3.1) in the previous section, except that s and f can depend upon r. Generally, both s and $s + f$ decrease in r.

The problem now has two dimensions. One is whether a borrower is eligible for a loan at all. The other problem is to determine the competitive rate when the borrower is eligible for a loan.

The fund manager's optimal *acceptance decision* is to lend to a borrower with $s[\cdot]$ and $f[\cdot]$ if and only if there exists an r satisfying (3.1) which, making the endogeneity of s and f explicit, can be rewritten as

$$(s[r] + \nu f[r])(1 + r) \geq 1 + \underline{r}. \tag{4.1}$$

Assuming that for an extraordinarily high rate r the probabilities $s[r]$ and $f[r]$converge to zero *faster than* r, that is

$$\lim_{r \nearrow \infty} (s[r] + f[r])(1 + r) \to 0,$$

the left-hand side of (4.1) can be plotted against r as in Figure 2. Namely, as in section 3, borrowers can be classified into the following three categories.

X: The set of those borrowers with $s[\cdot], f[\cdot]$ such that there exists an r satisfying $(s[r] + \nu_0 f[r])(1 + r) \geq 1 + \underline{r}$.

Y: The set of those borrowers with $s[\cdot], f[\cdot]$ such that there exists an r satisfying $(s[r] + f[r])(1 + r) \geq 1 + \underline{r}$, but that there exists no r satisfying $(s[r] + \nu_0 f[r])(1 + r) \geq 1 + \underline{r}$.

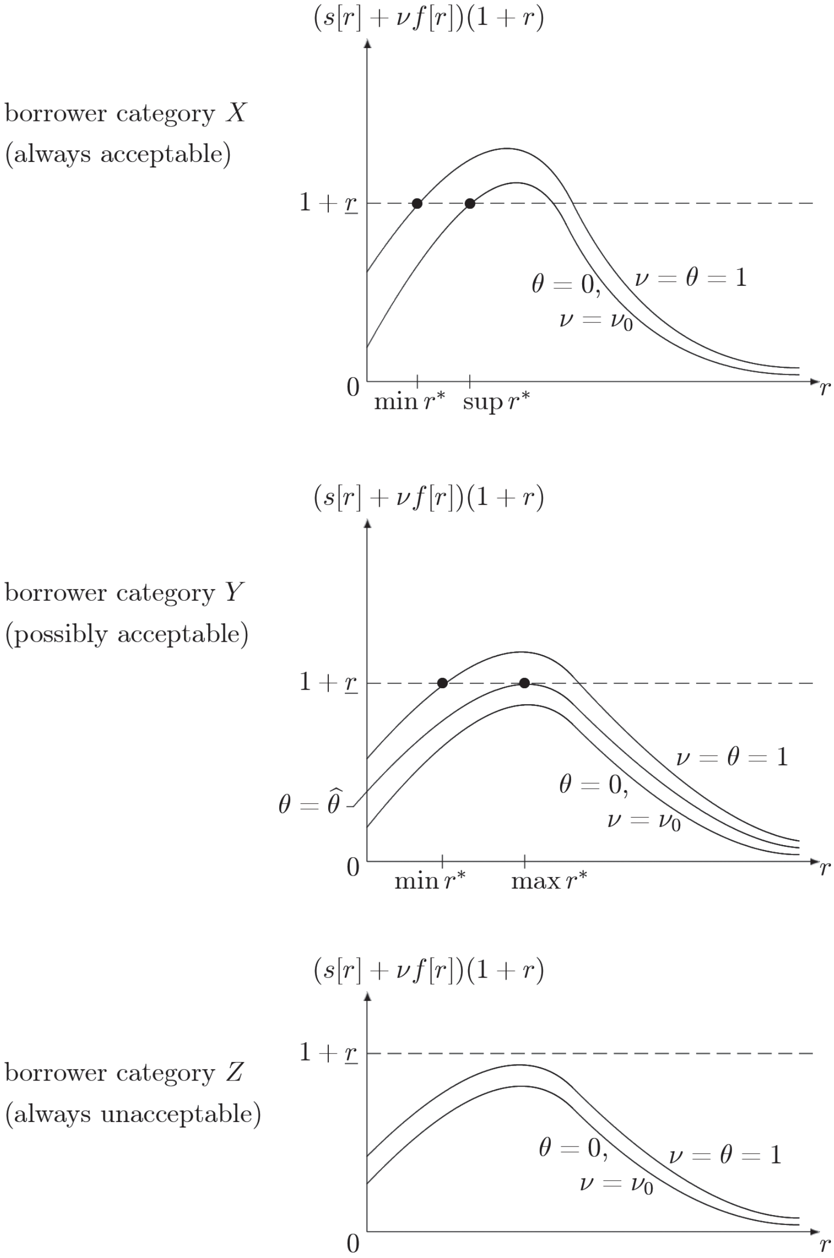
The lowest ν with which there exists an r satisfying $(s[r] + f[r])(1 + r) \geq 1 + \underline{r}$ is denoted as $\hat{\nu}$, and the fraction of fund managers accepting the borrower which entails this conditional probability $\hat{\nu}$ is henceforth denoted by $\hat{\theta}$.

Z: The set of those borrowers with $s[\cdot], f[\cdot]$ such that there exists no r satisfying $(s[r] + f[r])(1 + r) \geq 1 + \underline{r}$.

For borrower types X and Y, if there exists an r satisfying (4.1), then *the lowest value* of r, denoted by r^* hereinafter, defines the *competitive rate* for the borrower. Therefore, if everything is continuous, r^* should satisfy

$$(s\,[r^*] + \nu f\,[r^*])\;(1 + r^*) \geq 1 + \underline{r}. \tag{4.2}$$

Figure2: borrower types and ranges of competitive rates.



No competitive fund manager should accept any borrower in category Z. Also, fund managers can accept a borrower in category Y only if the borrower can expect a sufficiently high ν, which is made possibly only when a sufficiently high fraction of other fund managers are also ready to accept (i.e., refinance) the borrower. This is analogous to Proposition 1 in section 3.

A *major* difference here from the previous section is the competitiveness of the lending rates. In a competitive equilibrium, lenders are kept *indifferent* between trading with a qualified borrower (who satisfies (4.1)) and opting out for the safe alternative investment which earns $\underline{r}$ with certainty. This implies that, even when a borrower in either category X or category Y satisfies (4.1), a rational profit-maximising lender can still exercise the "discretion" not to accept the borrower. This implies the following.

Proposition 2 : It is a competitive equilibrium if, and only if, the fraction of fund managers accepting a borrower with $s[\cdot], f[\cdot]$ satisfies

$$\theta \begin{cases} = 0 & \text{for any} \quad (s[\cdot], f[\cdot]) \in Z, \\ \in [\hat{\theta}, 1] & \text{for any} \quad (s[\cdot], f[\cdot]) \in Y, \\ \in [0, 1] & \text{for any} \quad (s[\cdot], f[\cdot]) \in X, \end{cases}$$

and the competitive rate for any accepted borrower is determined as

$$r^* = \min \{ r \mid (s[r] + \nu f[r])(1 + r) \geq 1 + \underline{r} \} .$$

Compared with Proposition 1 and Corollary 1, endogenisation of lending rates can *enlarge* the possibility of discriminatory lending denial. Previously, only those borrowers in category Y were susceptible to possibly discriminatory "discretion" by fund managers. Now, not only category Y but also category X, economically the most "able" borrowers, can be subjected to fund managers' choosiness.

Furthermore, the following is particularly noteworthy.

Corollary 2 : For any borrower in either category X or category Y, the higher the fraction of accepting fund managers, the lower the resulting competitive rates.

For instance, if class-W borrowers in category X are accepted by all fund managers whilst class-B borrowers in category X are accepted by only a relatively small fraction of fund managers, then even between a class-W and class-B borrowers with the same $(s[\cdot], f[\cdot]) \in X$, the competitive rate for the class-W borrower is lower than that for the class-B borrower.

5. Imperfect competition

It is intriguing to discover that some aspects of competition are contributing to credit market discrimination. Competitiveness can materialise, and be conceptualised, in more than one way. Accordingly, the impact on credit market discrimination varies depending upon how the market is made imperfectly competitive.

Long-term lending relations

As aforementioned, establishing long-term lending relations can not only reduce the room for discrimination but more generally enhance intertemporal efficiency of lending and refinancing.

Realistically this may, in part, account for the “historical” background for credit discrimination. Namely, those borrowers who have had established long-term relations with lenders and fund managers may enjoy preferential treatment whilst those who has less established history, such as ethnic minorities consisting mostly of relatively recent immigrants, may face difficulties, *provided all their economic credentials being equal and all lenders and fund managers being economically rational*.

Supernormal profits

One of the most common ways to quantify competitiveness is the presence or the absence of supernormal profits. In this sense, competitiveness of the credit market implies that lenders cannot expect to earn supernormal profits. As shown in the previous section, this can have a very direct effect of making lenders indifferent between accepting and not accepting a perfectly “acceptable” borrower. If the credit market were slightly less competitive, so that lenders could expect small but nonnegligible supernormal profits from lending, then they would no longer reject a borrower "without reason". This not only eliminates the vast multiplicity of competitive equilibria in Proposition 2 but in particular, any $\theta \in [0, 1)$ for category X and $\theta \in [\hat{\theta}, 1)$ for category Y are eliminated, whereby the set of competitive equilibria will resemble that in Proposition 1 and Corollary 1.

Collusive lenders

In lieu of supernormal profits, imperfect competition may alternatively be defined as each individual lender and /or fund manager being no longer "small". Namely, if a nonnegligible mass of fund managers can collude, they may be able to affect ν .

It is intuitively obvious that collusion in the direction of *reducing* ν is never profitable. This eliminates, once again, any $\theta \in [0, 1)$ for category X and $\theta \in [\hat{\theta}, 1)$ for category Y. Furthermore, even when for some borrowers in category Y the *status quo* is $\theta = 0$, if more than $\hat{\theta}$ of the fund managers can collude, then they can "jump" the fraction from 0 to $\hat{\theta}$ and thenceforth the fraction can "trickle up" to 1.

6. Policy implications

Even though imperfect competition can eliminate some of the multiplicity of equilibria and thereby can serve as an anti-discrimination device to some extent, it does not necessarily follow that the policy maker should therefore strive to make the credit market as uncompetitive as possible. For, imperfect competition could often be accompanied by obvious negative side-effects, which might easily outweigh the possibly positive effects of anti-discrimination.

Before concluding this paper, it is worthwhile to investigate the possible effects of directly prohibiting discrimination based upon classes[3] . It might spontaneously appear as if the result of such a "class-blindness" policy could entail an outcome that is a population-weighted average of the previously equilibrium treatment of

3) Such prohibition has been implemented, for instance, in automobile insurance and health insurance where insurance providers are generally not allowed to collect information about the race or ethnicity of the insured.

each class of borrowers. Should this prediction be true, then if the population mass of the discriminated-against class is small relative to that of the discriminated-for class, the regulated outcome would vastly improve the treatment of the previously discriminated-against class in exchange for a relatively small sacrifice the previously discriminated-for class ought to endure. From a “democratic” point of view, such an outcome would be highly agreeable.

Upon a closer inspection, however, this spontaneous prediction may not always make sense. Return to the model presented in section 4. Suppose that there is a subset of Y where only class-W borrowers are currently offered loans whilst class-B borrowers are rejected. Now consider imposing a class-blind policy requiring the same treatment for both classes of borrowers whenever their types $s[\cdot]$ and $f[\cdot]$ are identical. The result from this policy is largely a matter of coordination. One of the possible ways to illustrate this situation is as follows.

1. If a fund manager conjectures that all other fund managers accept these borrowers in question, and hence he/she decides to accept them as well, then :
    - if the conjecture turns out correct, then the fund manager earns a normal profit from trading with these borrowers ;
    - however, if the conjecture turns out incorrect and few other fund managers in fact accept these borrowers, then he/she suffers a subnormal profit from trading with these borrowers.
2. If a fund manager conjectures that no other fund managers accept these borrowers in question, then he/she decides to decline these borrowers and hence will invest the fund in the safe alternative, unconditionally earning a normal profit.

Hence by weak dominance, the adjustment force one-sidedly favours not trading with these borrowers.

Obviously this is not the only possible way to describe the adjustment process : the outcome may depend upon the relation between the speed of adjustment in acceptance rules and that in competitive interest rates. A more detailed investigation in this direction can be a subject for future research. Nevertheless, the above suffices to illustrate in the most intuitive possible way why the seemingly no-nonsense anti-discrimination regulation might not bring about the desired effects.

7. Conclusion

By means of a relatively informal simple model, this paper has demonstrated that credit discrimination based upon *externally recognisable* yet *inherently non-economic* classes can be sustained as a self-fulfilling prophecy even *without* hidden actions taken by the borrowers. This differs in spirit from much of the standard theoretical explanation based upon the "vicious circle" where the borrowers are mistreated and hence are forced into risky ways of running the borrowed fund, resulting in high default rates which in turn confirms the lenders' expectation about these discriminated-against borrowers. The traditional vicious circle story could, depending upon how to interpret it, be used to "blame the victim": the theory gives little clue whether the chicken was first or the egg was first (therefore theoretically, it is never clear whether the victim is really a victim or merely an accomplice). In this paper, on the contrary, it is shown that discrimination can stand alone whilst the victim is entirely blameless (and hence is unquestionably a victim never an accomplice).

As a final note, it may take an extra care to apply this theory empirically. For, empirical data are scarcely helpful in inferring the contemplated use of the borrowed fund when the loan has been denied. Factually, however, credit discrimination appears to be a real issue (see, e.g., Blanchflower *et al*, 1999) whilst moral hazard, i.e., the "hidden" use of the borrowed funds by discriminated-against borrowers, has collected scarce evidence.

Bibliography

Arrow, K.J., "The Theory of Discrimination", in Ashenfelter, O. and Ree, A., eds., *Discrimination in Labor Markets* (1973), Princeton University Press.

Blanchflower, D.G., Levine, P.B. and Zimmerman, D.J., "Discrimination in the Small Business Credit Market" (1999), Dartmouth College Working Paper 99-08.

Coate, S. and Loury, G., "Will Affirmative-Action Policies Eliminate Negative Stereotypes?", *American Economic Review*, Vol. 83 (1993), pp. 1220-40.