Stereotypes and Inequality: A 'Signaling' Theory of Identity Choice

Young Chul Kim
(Associate Research Fellow, Korea Development Institute)

Glenn C. Loury
(The Merton P. Stoltz Professor of the Social Sciences and
Professor of Economics, Brown University)

고정관념과 불평등: 정체성 선택에 관한 신호이론

김 영 철
(한국개발연구원 부연구위원)

글렌 라우리
(브라운대학교 멀튼 스톨츠 사회과학 석좌교수 겸 경제학과 정교수)

* 김영철: (e-mail) yckim@kdi.re.kr, (address) Korea Development Institute, Hoegiro 47, Dongdaemun-gu Seoul, Korea.
글렌 라우리: (e-mail) Glenn_Loury@brown.edu, (address) Department of Economics, Brown University, Providence, RI 02912, USA.

ABSTRACT

We develop an identity choice model within the context of a stereotyping-cum-signaling framework. The model allows us to explore implications of the fact that, when individuals can choose identity, then the distribution of abilities within distinct identity groups becomes endogenous. This is significant because, when identity is exogenous and if the ability distributions within groups are the same, then inequality of group reputations in equilibrium can only arise if there is a positive feedback between group reputation and individual human capital investment activities (Arrow, 1973; Coate and Loury, 1993). Here we show that when group membership is endogenous then the logic of individuals' identity choices leads there to be a positive selection of higher ability individuals into the group with a better reputation. This happens because those for whom human-capital-investment is less costly are also those who stand to gain more from joining the favored group. As a result, ability distributions within distinct groups can endogenously diverge, reinforcing incentive-feedbacks. We develop the theoretical framework that can examine the positive selection and the endogenous group formation. The model implies that inequality deriving from stereotyping of endogenously constructed social groups is at least as great as the inequality that can emerge between exogenously given groups.

　본 논문은 고정관념(stereotypes)과 노동시장 신호(signaling)에 관한 이론적 틀에 기반하여 개인의 정체성(identity) 선택에 관한 이론모형을 제시하고 있다. 사회집단 간의 평판의 차이 및 이로 인한 집단 간 불평등은 사회집단에 대한 시장의 고정관념 및 이에 상응하는 집단 구성원의 인적자본 투자수준에 의해 형성된다(Arrow[1973]; Coate and Loury[1993]). 한편, 이제까지의 학계의 논의는 개인들의 사회집단 선택 및 이에 따른 사회집단 구성의 내생적 변화 가능성을 배제한 채 발전하여 왔다. 사회집단 간 평판의 차이가 존재하고 동시에 개인들의 자발적 집단 선택이 가능한 상황이라면, 평판이 좋은 사회집단은 집단의 구성원들이 보다 적극적으로 인적자본에 투자하는 '투자 확대'의 효과뿐만 아니라, 평판이 상대적으로 떨어지는 사회집단에 비해 보다 우수한 인재들을 집단의 새로운 구성원으로 맞아들이는 '집단 구성'의 효과를 동시에 누리게 된다. 이러한 집단 구성의 내생성은 본질적으로 사회집단 간의 평판의 차이를 보다 확대시키는 결과를 낳게 되며,

ABSTRACT

이렇게 확대된 평판의 격차가 우수한 인재의 집중을 보다 강화시킴으로써 집단 간 경제적 불평등이 시간이 지남에 따라 악화되는 것이다. 본 논문은 노동시장 신호이론을 활용하여 사회집단 선택의 내생성하에서의 집단 간 불평등의 균형점을 찾고 이들의 안정성에 관해 논하고 있으며, 집단 구성의 내생성이 주어질 경우 그렇지 않은 경우에 비해 사회집단 간의 경제사회적 불평등이 보다 심화될 수 있음을 이론적으로 증명하고 있다.

Ⅰ. Introduction

We develop an identity choice model based on a stereotyping-cum-signaling framework. If a worker's productivity is not perfectly observable, employers have an incentive to use the collective reputations of the identity groups to which job applicants belong in the screening process. A key feature of our model is that individuals belonging to a group with a good collective reputation have a greater incentive to invest in skills than do those who belong to a group with a poor reputation. And yet, given its greater rate of investment in skills, the former group will tend to maintain a better collective reputation than the latter. That is, there can be more than one self-confirming equilibrium in group reputations (Arrow, 1973; Coate and Loury, 1993). Previous work on such statistical discrimination has taken group identity as immutable, so that each individual is affected by the collective reputation of his own group only. We handle the dynamics between the collective reputation and the identity choice problem by relaxing this immutability assumption.

To illustrate the relevance of our model, consider one concrete example—the phenomenon commonly known as "passing." This widely observed behavior is evidently an instance of identity choice. Talented young members in the group with a worse collective reputation may consider "passing" into the group with a better reputation when the return for such "passing" (e.g., better treatment in the labor market) outweighs its cost (e.g., loss of ties to one's own kind.) Thus, it has been noted that a significant number of the blacks in the U.S. population consistently passes for White or some other race (Sweet, 2005).[1] Moreover, many ethnic Koreans in Japan (most of whom descend from forced laborers in mines and factories who were brought to Japan from the Korean peninsula during the period of Japanese imperialism) are "passing" for native Japanese—by changing both their surnames and their given names when seeking formal employment or marriage. They have done so in order to escape negative stereotypes and prejudices against ethnic Koreans in Japan.[2]

When passing for a member of the advantaged group with high reputation is not possible due to immutability, the most talented of the stereotyped group are more likely to seek styles of self-presentation that aim to communicate "I'm not one of THEM; I'm one of YOU!" because they are the ones who gain most by separating themselves from the mass (Loury, 2002). Taking the example of the Black population in the US, methods that are known to be used for "partial passing" are: affections of speech, dressing up rather than wearing casual clothes, spending more on conspicuous consumption, and migration to affluent residential areas (Charles *et al.*,

---

[1] The National Longitudinal Survey conducted by the Department of Labor of the US shows that 1.87 percent of those who had originally answered "Black" to the interviewer's race question in 1979 switched to either "White," "I don't know, " or "other," by 1998.

[2] Every year about 10,000 Koreans living in Japan, out of around 600,000 Korean descendants holding Korean nationality, choose to be naturalized, giving up their names and original nationality (Fukuoka *et al.*, 1998).

2009; Goffman, 1959). There is evidence that the more educated (or talented) blacks tend to speak standard American English rather than African American English (Grogger, 2011). That is, the most talented of the stereotyped group "pass for" the slightly better-off subgroup that maintains a higher reputation than the stereotyped population as a whole by adopting the cultural traits of the subgroup.

This selective out-migration to the better-off subgroup may undermine solidarity in the disadvantaged population and cause conflict among them, such as the accusation of "Acting White" against the ones who practice the "partial passing" methods (Fryer and Torelli, 2010). However, there might be a social gain through this practice: at least some cultural subgroups of the stereotyped population might be able to recover their reputation when the talented young members gather around certain cultural traits. The usage of the observable cultural traits in the screening process can to some extent cure the social inefficiency caused by imperfect information about the true characteristics of workers.

It is not only backward or disadvantaged groups whose behavior is captured by our model. The emergence of an elite social group out of a population can also be explained through identity choice behavior. Fang (2001) discussed the use of cultural instruments that are intrinsically irrelevant for productivity to form an elite group. He notes that this may help to an account for the complexity of elite etiquette in European (or Confucian) societies, as well as for the greater respect afforded to those with an "Oxford Accent." Skilled and unskilled workers have different incentives to join a group with unique cultural traits that are expensive to obtain. Thus, the cultural group is treated preferentially by employers due to the higher fraction of skilled workers, even though the cultural traits of the group are not relevant for productivity. We may see an autonomously growing elite subgroup with differentiated cultural traits whose members are considered as distinguished from their peers.

The identity choice model in this paper starts with a standard statistical discrimination framework (Coate and Loury, 1993). We identify multiple self-confirming prior beliefs, which we call Phenotypic Stereotyping Equilibria (PSE). This multiplicity of phenotypic equilibrium explains inequality of collective reputations between exogenous and equally endowed identity groups as being due to positive feedback between a group's reputation and its members' investment incentives. But it entails no selection into or out of the groups. However, when membership is endogenous and if the groups' reputations were to differ in equilibrium, then not only will members of a favored group face greater incentives to invest in human capital. It is also the case that the better-regarded group will, in equilibrium, come to consist disproportionately of high ability/low human capital investment cost types. These are the types who gain most from joining a favored group. The result is that human capital cost distributions between groups endogenously diverge, which reinforces incentive-feedbacks. We call such a group-disparate equilibrium with positive selection a (non-trivial) Affective Stereotyping Equilibrium (ASE).

For the development of our theoretical model, we introduce two affects (i.e., possible group identities), A and B. We assume that the cost to choose affect A rather than B varies across the population. Agents choose affect A if and only if the anticipated return exceeds the agent's cost of choosing affect A. We further assume

that the cost distribution among agents for this affect choice is statistically independent of the cost distribution of human capital investment. (That is, a person's identity orientation cannot be used to predict their economic abilities.) Nevertheless, we show that in the model's equilibrium the more able individuals—those with lower human capital investment costs—will tend to choose affect A whenever the collective reputation of the affect A group is better than that of the affect B group. But, of course, their behaving in this way is what causes group A to have a better reputation in the first place!

The main result which we demonstrate with this model is the following: group inequality that derives from the ascriptive stereotyping of endogenously constructed social groups is at least as great as the group inequality that can emerge from the phenotypic stereotyping of exogenously given groups. Again, this is due to the fact that when groups are endogenous then low human capital cost types are disproportionately drawn to the group with a better reputation, causing a skill disparity between groups to endogenously diverge, thereby justifying the disparity of group reputations. The model also implies that there exist multiple (non-trivial) Affective Stereotyping Equilibria whenever multiple Phenotypic Stereotyping Equilibria can be constructed in the labor market structure.

The paper is organized into the following sections. Section 2 develops the model with the identify choice and skill investments. Section 3 defines Phenotypic and Affective Stereotyping Equilibria. Section 4 discusses the implications of the developed model and Section 5 concludes.

II. Model with the Identity Choice

In this section, we display the general framework of the model that includes agents' decision making and the payoff structure. The workers make an investment decision on skill acquisition and choose the identity type before they enter the labor market. Employers set the wage for each worker to be proportional to the worker's expected productivity using both the productivity-related noisy signal and the identity type. The workers' decision makings, the employers' wage setting and the consequent expected payoff are discussed below.

Workers' Affective/Expressive Behavior: Agents choose affect $i \in \{A, B\}$. The cost to choose the affect A is $k \in R$. k can be negative: the affect A can generate benefits for some agents. CDF of the affective behavior cost is denoted by $H(k)$. We assume the affective symmetry: $H(k) = 1 - H(-k)$. Agents choose the affect A if and only if the anticipated return exceeds the agent's cost k. Otherwise, they choose the affect B. WLOG, it is natural to assume that PDF of the cost k, $h(k)$, has one peak at $k=0$: $h'(k)>0$ for any $k \in (-\infty, 0)$ and $h'(k)< 0$ for any $k \in (0, \infty)$.[3]

Workers' Skill Acquisition Behavior: Agents choose whether to be skilled or not: $e \in \{0, 1\}$. The cost to be skilled is c, which is non-negative. CDF of the skill

---

[3] This is not a critical assumption in the model. We will have the identical conclusions with the uniform distribution of k.

acquisition cost is $G(c)$, in which $G(0) \geq 0$ and $G(\infty) = 1$.[4] WLOG, it is natural to assume that PDF of the cost c, $g(c)$, has one peak at $\hat{c}$: $g'(c) > 0$ for any $c \in (0, \hat{c})$ and $g'(c) < 0$ for any $c \in (\hat{c}, \infty)$. An agent chooses ($e = 1$) if the return from doing so exceeds that agent's cost for the skill acquisition (c). We impose that c and k are independently distributed.

Employers' Wage-setting Behavior: Skill e is not fully identified. Employers observe group identity and noisy signal $t \in R^+$ distributed conditional on e. PDF of the signal conditional on e is $f_e(t)$ and its CDF is $F_e(t)$. Let us define the function $f(\pi, t)$ as $f(\pi, t) \equiv \pi f_1(t) + (1 - \pi) f_0(t)$, which indicates the distribution of signal t of agents belonging to a group of which the skill level(the fraction of the skilled workers) is believed to be π. WLOG, we assume that $\frac{f_1(t)}{f_0(t)}$ increases with respect to t, which is denoted by MLRP: Monotonic Likelihood Ratio Property. The employers' belief that an agent with signal t is skilled is $\rho(\pi, t) (\equiv Pr[e = 1 | \pi, t]) = \frac{\pi f_1(t)}{f(\pi, t)}$. Under MLRP, $\rho(\pi, t)$ increases with both π and t. The productivity of a skilled worker is ω and that of an unskilled worker is zero. We assume that the wage is proportional to the expected skill level:

$$W(\pi, t) = \omega \cdot \rho(\pi, t), \text{for some } \omega > 0$$

$$= \omega \cdot \frac{\pi f_1(t)}{\pi f_1(t) + (1 - \pi) f_0(t)}. \tag{1}$$

Workers' Payoffs: The expected wage from acquiring skill level e is denoted by $V_e(\pi)$:

$$V_e(\pi) = \int_0^1 f_e(t) W(\pi, t) \, dt, \tag{2}$$

in which $V_e'(\pi)$ is positive for any $e \in \{0,1\}$. Workers' expected return acquiring human capital $(R(\pi))$ is defined as

$$R(\pi) \equiv V_1(\pi) - V_0(\pi). \tag{3}$$

$R(\pi)$ is expressed as

$$R(\pi) = \int_0^1 (f_1(t) - f_0(t)) W(\pi, t) \, dt$$

$$= \omega \pi \int_0^1 \frac{(f_1(t) - f_0(t)) f_1(t)}{f(\pi, t)} \, dt. \tag{4}$$

The followings can be easily seen

---

[4] With $G(c) \geq 0$, we allow that a fraction of workers always invest for skills.

$$R'(\pi) = \omega \int_0^1 \frac{(f_1(t) - f_0(t))f_1(t)f_0(t)}{f(\pi,t)^2} dt, \tag{5}$$

$$R'(\pi) = -2\omega \int_0^1 \frac{(f_1(t) - f_0(t))^2 f_1(t)f_0(t)}{f(\pi,t)^3} dt \ (< 0). \tag{6}$$

Thus, $R(\pi)$ is concave and $R(0) = R(1) = 0$, which implies that $\lim_{\pi \to 0} R'(\pi) > 0$ and $\lim_{\pi \to 1} R'(\pi) < 0$. Let us denote $argmax\{R(\pi)\}$ by $\bar{\pi}: R'(\bar{\pi}) = 0$.

The first derivatives of $V_0(\pi)$ and $V_1(\pi)$ are

$$V_0'(\pi) = \int_0^1 \omega f_1(t)f_0(t)^2 f(\pi,t)^{-2} dt, \tag{7}$$

$$V_1'(\pi) = \int_0^1 \omega f_1(t)^2 f_0(t) f(\pi,t)^{-2} dt. \tag{8}$$

Note that $V_0'(0) = \omega$ and $V_1'(1) = \omega$. Since we know $R'(0) > 0$ and $R'(1) < 0$, we have $V_0'(1) > \omega$ and $V_1'(0) > \omega$. It is more likely that $V_0'(\pi)$ tends to increase as π increases and $V_1'(\pi)$ tends to decrease as π increases.

Thus, a worker with cost c, in a group believed to be investing at rate π has the payoff:

$$U(\pi,c) = \max\{V_1(\pi) - c; \ V_0(\pi)\}, \tag{9}$$

in which the function $U(\pi,c)$ is increasing in $\pi (\because V_e'(\pi) > 0, \forall e \in \{0,1\})$ and non-increasing in c.

Ⅲ. Stereotyping Equilibrium

In this section, we define the Phenotypic and Affective Stereotyping Equilibria and search for the properties of the equilibria.

1. Phenotypic Stereotyping Equilibria

Given the employers' prior belief (π) about human capital investment rate in a population, the fraction of workers who choose $(e=1)$ is $G(R(\pi))$. Let us denote an equilibrium belief/investment rate by $\hat{\pi} \in [0,1]: \hat{\pi} = G(R(\hat{\pi}))$. The set of all such equilibria is denoted by Ω_{CL} (Coate and Loury, 1993). Let us call them Phenotypic Stereotyping Equilibria (PSE). Inequality of collective reputation between exogenous groups in equilibrium is due to feedback between group reputation and individual investment activities. The individuals in a group with a better collective reputation have a greater incentive to invest in skills, and with their greater skill investment rate, the group maintains a better collective reputation, (and Vice Versa).

It is most likely that there exists either one or three equilibria in the economy, because $G(c)$ is S-shaped as displayed in Figure 1. Multiple equilibria $\hat{\pi} \in \Omega_{CL}$ create possibility of Phenotypic Stereotyping (PS) wherein groups are exogenously and visibly distinct, though equally well endowed. Nevertheless, they fare unequally in the equilibrium (Panel B of Figure 1).

It is clear that the socially optimal level of human capital investment is $G(\omega)$. However, human capital investment is socially inadequate in any PSE $\hat{\pi}: G\big(R(\hat{\pi})\big) < G(\omega)$ for any $\hat{\pi} \in \Omega_{CL}$, because $R(\hat{\pi}) < \omega$ as $\omega \int_0^1 f_1(t)dt - R(\pi) = \omega \int_0^1 f_1(t) f_0(t) f(\pi,t)^{-1} dt > 0$. This implies the social optimality can be further improved through the relaxation of the given constraints.

2. Affective Stereotyping Equilibria

Absent affective discrimination, workers choose 'affect' based on their "natural" orientation: $i = B$ if $k > 0$ and $i = A$ if $k < 0$. This implies that the human capital cost distribution, namely $G(c)$, is the same for both affective groups. Refer to Figure 2 for the case with affective discrimination absent. However, when affective discrimination prevails in the labor market and group membership is endogenous, this is not true any more. Given that the initial reputations of the two affective groups differ, the favored group not only faces great human capital investment incentives, but it also consists disproportionately of low human capital investment cost types, who gain more from joining a favored group. Thereby, it will cause human capital cost distributions between groups to endogenously diverge, reinforcing incentive-feedbacks.

The economic analysis of the story can be developed in the following way. Let π_i be employer belief about human capital investment rate in affective group i. Consider two affective groups A and B. Let us define a function $\Delta U(\pi_A, \pi_B; c)$ as the payoff difference between a A-type worker and a B-type worker given their skill acquisition cost level c: $\Delta U(\pi_A, \pi_B; c) \equiv U(\pi_A, c) - U(\pi_B, c)$. Given $\pi_A > \pi_B$, $\Delta U(\pi_A, \pi_B; c)$ is positive because $\partial U(\pi, c)/\partial \pi > 0$. Note that $\Delta U(\pi_A, \pi_B; c) = -\Delta U(\pi_B, \pi_A; c)$ and $\Delta U(\pi, \pi; c) = 0$.

An agent with the cost set (c, k) chooses affective behavior $i = A$ if and only if $\Delta U(\pi_A, \pi_B; c) \geq k$. Otherwise, he chooses affective behavior $i = B$. Given that c and k are independent, the fraction of agents choosing $(i = A)$ is given by

$$\Sigma^A \equiv \int_0^\infty H(\Delta U(\pi_A, \pi_B; c)) \, dG(c). \tag{10}$$

The fraction of workers choosing $(i = A)$ and $(e = 1)$ is given by

$$\sigma^A \equiv \int_0^{R(\pi_A)} H(\Delta U(\pi_A, \pi_B; c)) \, dG(c). \tag{11}$$

Then, the fraction of agents choosing $(i = B)$ is obtained using $\Sigma^B = 1 - \Sigma^A$ and $\Delta U(\pi_A, \pi_B; c) = -\Delta U(\pi_B, \pi_A; c)$:

$$\Sigma^B \equiv \int_0^\infty H(\Delta U(\pi_B, \pi_A; c)) \, dG(c). \tag{12}$$

Consequently, the fraction of workers choosing ($i=B$) and ($e=1$) is given by

$$\sigma^B \equiv \int_0^{R(\pi_B)} H(\Delta U(\pi_B, \pi_A; c)) \, dG(c). \tag{13}$$

Given the employer belief about human capital investment rates (π_A, π_B), the actual investment rates for the affective groups denoted by $\phi(\pi_A, \pi_B)$ and $\phi(\pi_B, \pi_A)$ are

$$\begin{cases} Pr\{e=1 | i=A, \pi_A, \pi_B\}(\equiv \phi(\pi_A, \pi_B)) = \sigma^A/\Sigma^A, \\ Pr\{e=1 | i=B, \pi_B, \pi_A\}(\equiv \phi(\pi_B, \pi_A)) = \sigma^B/\Sigma^B. \end{cases} \tag{14}$$

It is noteworthy that when employers' belief is the same for both affective groups $(\pi_A = \pi_B)$, $\Delta U(\pi_B, \pi_A; c)$ is zero and we have $R(\pi_A) = R(\pi_B)$. This implies that the affective behavior does not affect the human capital investment activities: $\phi(\pi_A, \pi_B) = \phi(\pi_B, \pi_A)(= G(R(\pi_A)))$.

An equilibrium with affective stereotyping (ASE) is defined as a pair of investment rates for the affective groups $(\pi_A^*, \pi_B^*) \in [0, 1]^2$ such that $\pi_A^* = \phi(\pi_A^*, \pi_B^*)$ and $\pi_B^* = \phi(\pi_B^*, \pi_A^*)$. The set of all such equilibria is denoted by Ω_{KL}. Note that every PSE corresponds to trivial ASE where differences in affect are uninformative: $(\hat{x}, \hat{x}) \in \Omega_{KL}$ if $\hat{x} \in \Omega_{CL}$ because $\phi(\hat{x}, \hat{x}) = G(R(\hat{\pi})) = \hat{x}$. Affective stereotyping discrimination occurs if and only if $\pi_A^* \neq \pi_B^*$.

Note that $\Delta U(\pi_A, \pi_B; c)$ can be expressed by

$$\Delta U(\pi_A, \pi_B; c) = \max\{R(\pi_A) - c; 0\} + V_0(\pi_A) - \max\{R(\pi_B) - c; 0\} - V_0(\pi_B). \tag{15}$$

Using $R(\pi_A)$ and $R(\pi_B)$, we have the following lemma concerning $\Delta U(\pi_A, \pi_B; c)$:

Lemma 1. *For any* $c \leq m\dot{n}\{R(\pi_A), R(\pi_B)\}, \Delta U(\pi_A, \pi_B; c) = V_1(\pi_A) - V_1(\pi_B)$. *For any* $c \geq m\,ax\{R(\pi_A), R(\pi_B)\}, \Delta U(\pi_A, \pi_B; c) = V_0(\pi_A) - V_0(\pi_B)$. *For any* *c* *such that* $m\dot{n}\{R(\pi_A), R(\pi_B)\} < c < m\,ax\{R(\pi_A), R(\pi_B)\}$, *we have*

$$\Delta U(\pi_A, \pi_B; c) = \begin{cases} V_1(\pi_A) - V_0(\pi_B) - c & f \quad R(\pi_A) \geq R(\pi_B), \\ V_0(\pi_A) - V_1(\pi_B) + c & f \quad R(\pi_A) < R(\pi_B). \end{cases} \tag{16}$$

The above lemma is summarized in Figure 3. Panel A of the figure displays the case with $\pi_A > \pi_B$ and panel B does the case with $\pi_A < \pi_B$. It is easily seen that $\Delta U(\pi_A, \pi_B; c) > 0$ for any *c* if and only if $\pi_A > \pi_B$. Therefore, we have the following result.

Proposition 1. *When employers have different beliefs about two affective groups* $(\pi_A \neq \pi_B)$,

the number of workers who adopt the 'affect' corresponding to the favored employers' belief is greater than that of workers who adopt the 'affect' with the less favored employers' belief: $\Sigma^i > \Sigma^j$ if $\pi_i > \pi_j$ for any $i, j \in \{A, B\}$.

That is, in the current setting with symmetric cost distribution, more than half workers adopt the 'affect' that corresponds to the more favorable employers' belief: $\Sigma^i > .5$ and $\Sigma^j < .5$ if $\pi_i > \pi_j$ for any $i, j \in \{A, B\}$. The Lemma 1 implies that $\Delta U(\pi_A, \pi_B; c)$ is non-increasing with respect to c whenever $R(\pi_A) > R(\pi_B)$, and non-decreasing whenever $R(\pi_B) > R(\pi_A)$. It leads to the following useful lemma.

Lemma 2. *Whenever $R(\pi_A) > R(\pi_B)$, the following holds: $\phi(\pi_A, \pi_B) > \phi(\pi_A, \pi_A)$ and $\phi(\pi_B, \pi_A) < \phi(\pi_B, \pi_B)$. In a symmetric way, whenever $R(\pi_A) < R(\pi_B)$, the following holds: $\phi(\pi_A, \pi_B) < \phi(\pi_A, \pi_A)$ and $\phi(\pi_B, \pi_A) > \phi(\pi_B, \pi_B)$. Even when $\pi_A \neq \pi_B$, the following holds if $R(\pi_A) = R(\pi_B)$: $\phi(\pi_A, \pi_B) = \phi(\pi_B, \pi_A) = \phi(\pi_A, \pi_A) = \phi(\pi_B, \pi_B)$.*

The above lemma implies the following proposition.

Proposition 2. *The disproportionately more talented workers, whose human capital investment costs (c) are relatively lower, choose the 'affect' that corresponds to the greater return to human capital investment: given $R(i) > R(j)$, $\phi(i, j) > G(R(i))$ and $\phi(j, i) < G(R(j))$ for any $i, j \in \{\pi_A, \pi_B\}$.*

Ⅳ. Implications and Further Discussion

It is obvious that (x, x) is an affective stereotyping equilibrium for any phenotypic stereotyping equilibrium x because $x = \phi(x, x)$. These ASE are called trivial ASE in this paper. We are more concerned with the non-trivial ASE in which the affective group's reputations differ in equilibria: $x \neq y$. Intuitively, we can say that the existence of non-trivial ASE is guaranteed given the multiple PSE. Suppose that there exist three PSE (π_h, π_m, π_l) in the given labor market structure as displayed in Panel B of Figure 1. Imagine that an affective group's reputation is the level of π_h and the other affective group's reputation is that of π_l. Note that the following must be true given the multiple PSE: $R(\pi_h) > R(\pi_l)$. Then, the combination (π_h, π_l) cannot be stable because of Proposition 2: given $R(\pi_h) > R(\pi_l)$, $\phi(\pi_h, \pi_l) > G(R(\pi_h))$ and $\phi(\pi_l, \pi_h) < G(R(\pi_l))$. That is, the skill investment rate of the entering cohort to the affective group with π_h is greater than the current skill level of the group $G(R(\pi_h))$, while that of the entering cohort to the other affective group with π_l is smaller than the current skill level of that group $G(R(\pi_l))$. The situation is well described in Panel A of Figure 3. Therefore, there must be at least two non-trivial ASE: one can emerge out of the initial combination (π_h, π_l) and the other out of the initial combination (π_l, π_h).

We also know that the skill difference between two entering cohorts $(\phi(\pi_h, \pi_l) - \phi(\pi_l, \pi_h))$ is greater than that between the two groups' current skill levels $(\pi_h - \pi_l)$. This implies that the inequality between endogenous groups in some non-trivial ASE can be greater than inequality between exogenous groups in any PSE.

Finally, to see how our model can be used to shed light on labor market phenomena other than ethnic/racial group inequality, let us consider the problem of "re-branding ex-cons." The main feature of the labor market for ex-convicts is that employers wish to avoid associating with those who end up returning to criminal activity, but employers cannot be certain from information which among the convicts will and which will not do so. However, the convicts themselves are presumed to know their own intentions. Under this informational asymmetry, employers choose not to hire any ex-cons leaving all of them unemployed.

Let us imagine a *re-branding program* run by the government along the following line: There is to be a certifiable and costly activity such that, before going into the labor market, ex-cons can elect to join this program or not. Those who will go straight are more willing to join the program than those who will return to crime because those going straight have greater gain from having a job in the market. The collective reputation of the ex-con subgroup with the program certificate improves with the greater fraction of ex-cons going straight joining the program, so that those with the program certificate will be hired and those without it will not be hired by employers in the labor market. That is, notwithstanding the informational asymmetry and the *adverse selection* issue in the market, a government can nevertheless design a costly program by means of which some ex-cons can credibly convey their good intentions to employers. The social efficiency can be achieved by the introduction of the costly program.

Ⅴ. Conclusion

In this paper, we develop an identity choice model that can explain social activities such as passing and selective out-migration from a stereotyped group, loosening the assumption of group identity immutability in standard statistical discrimination models. More talented members with low human capital investment cost have a greater incentive to identify themselves with a group that has a better collective reputation. The positive selection into a favored group plays a critical role in causing human capital cost distributions between groups to endogenously diverge. This model can be applied to many other social settings such as code switching (Goffman, 1959) and generating certificates to fight against negative stereotypes (e.g. re-branding ex-cons).

References

Arrow, K. J., "The Theory of Discrimination," in Ashenfelter and Rees (eds.), *Discrimination in Labor Markets*, 1973.

Charles, Kerwin K., Erik Hurst, and Nikolai Roussanov, "Conspicuous Consumption and Race," *The Quarterly Journal of Economics* 124, 2009, pp.425~467.

Coate, Stephen and Glenn Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotype?" *American Economic Review* 83, 1993, pp.1220~1240.

Fang, Haming, "Social Culture and Economic Performance," *American Economic Review* September, 2001, pp.924~937.

Fukuoka, Yasunori, Yukiko Tsujiyama, and John Russell, "MINTOHREN: Young Koreans Against Ethnic Discrimination in Japan," *The Bulletin of Chiba College of Health Science* 10, 1998.

Fryer, Roland and Paul Torelli, "An Empirical Analysis of 'Acting White'," *Journal of Public Economics* 94, 2010, pp.380~396.

Goffman, Erving, *The Presentation of Self in Everyday Life*, New York: Doubleday, 1959.

Grogger, Jerey, "Speech Patterns and Racial Wage Inequality," *The Journal of Human Resources* 46, 2011, pp.1~25.

Loury, Glenn C., *The Anatomy of Racial Inequality*, Harvard University Press, 2002.

Sweet, Frank W., "The Rate of Black-White 'Passing'," *Legal History of the Color Line: The Rise and Triumph of the One-Drop Rule*, Backintyme, 2005.
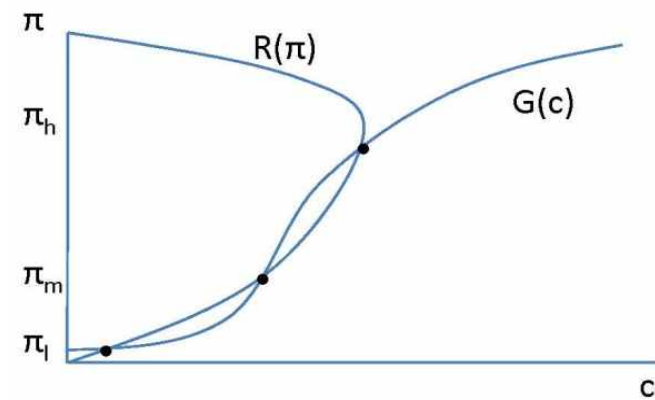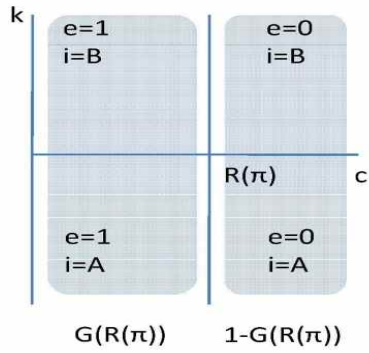
[Figure 1] Phenotypic Stereotyping Equilibria

Panel A. Unique PSE
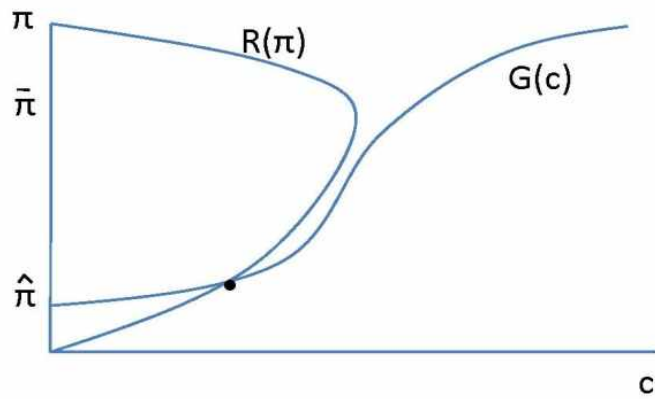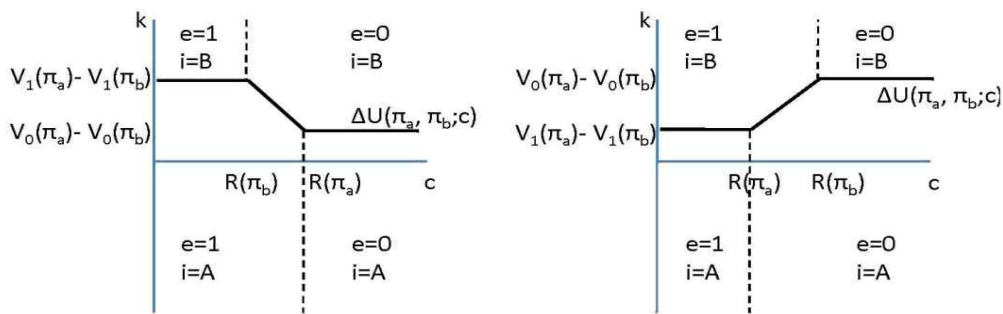

Panel B. Mulitple PSE

[Figure 2] Human Capital Investment with Affective Discrimination Absent

[Figure 3] Human Capital Investment with Affective Discrimination

Panel A. Case with $\pi_a > \pi_b$

Panel B. Case with $\pi_a < \pi_b$