

# AN ECONOMIC ANALYSIS OF ‘ACTING WHITE’\*

DAVID AUSTEN-SMITH AND ROLAND G. FRYER, JR.

February 28, 2005

## Abstract

This paper formalizes a widely discussed peer effect entitled ‘acting white’. ‘Acting White’ is modeled as a two audience signaling quandary: signals that induce high wages can be signals that induce peer group rejection. Without peer effects, equilibria involve all ability types choosing different levels of education. ‘Acting White’ alters the equilibrium dramatically: the (possibly empty) set of lowest ability individuals and the set of highest ability individuals continue to reveal their type through investments in education; ability types in the middle interval pool on a common education level. Only those in the lower intervals are accepted by the group. The model’s predictions fit many stylized facts in the anthropology and sociology literatures regarding social interactions among minority group members.

---

\*This paper is the melding of two independent projects: “Peer Pressure and Job Market Signaling” [Austen-Smith 2002] and “The Economics of Acting White” [Fryer 2001]. The formal results of these independent papers were essentially identical, thus, we decided to put them together. We are grateful to Roger Myerson for putting us in touch with one another. We are also grateful to Gary Becker, Phillip Cook, Matthias Dahm, Ronald Ferguson, Edward Glaeser, Caroline Hoxby, Lawrence Katz, Steven Levitt, Gavin Samms, Jesse Shapiro, two anonymous referees, and seminar participants too numerous to mention for comments and suggestions. This paper corrects Austen-Smith and Fryer [2003], and all previous versions, which omitted an equilibrium possibility. Financial support from the John D. and Catherine T. MacArthur Foundation through the Social Interactions and Inequality Network (Austen-Smith) and the National Science Foundation (Fryer: SES-0109196) is gratefully acknowledged. The usual caveat applies.

“Some African-American students, unable to extricate themselves from the quicksand of self-defeat, have adopted the incredibly stupid tactic of harassing fellow blacks who have the temerity to take their studies seriously. According to the poisonous logic of the harassers, any attempt at acquiring knowledge is a form of ‘acting white.’ ”

(Bob Herbert, *The New York Times*, March 1, 1995)

## I. INTRODUCTION

On every measure of academic achievement black students lag behind their white counterparts.<sup>1</sup> In 2000, National Association of Education Progress (NAEP) scores revealed that black 17 year olds read at the proficiency level of white 13 year olds. On the Scholastic Aptitude Test (SAT), the virtual gateway to America’s elite colleges and universities, there is little overlap in the *distribution* of black and white test scores. And, these differences tend to be exacerbated in highly segregated areas [Cutler and Glaeser 1997 and Case and Katz 1991].<sup>2</sup> Economists and sociologists often argue that these dramatic differences in performance are the result of negative peer interactions or spillovers – amorphous terms with little specific economic content. Gaining a better understanding of peer effects which contribute to black underachievement is of paramount importance in forming public policy and the subject of this paper.

Many ethnographers argue that peer effects take a particularly insidious form: black peers and communities impose costs on their members who try to ‘act white’ [Fordham and Ogbu 1986, Fordham 1996, Corwin 2001, Suskind 1998].<sup>3</sup> Individuals exposed to these social interactions have disincentives to invest in particular behaviors (i.e. education, ballet, proper speech) due to the fact that they may be rejected by their social peer group. This behavior is counter-intuitive as one might presume that communities have an incentive to push its members out and support their success. To be sure, this occurs within some ethnic groups (e.g. Chinese Americans and American Jews), but there are many groups across the world who face the dilemma between loyalty and success described here.<sup>4</sup>

This paper formalizes a particular peer effect, ‘acting white,’ which potentially contributes to

---

<sup>1</sup>This fact was first established, systematically, in the Coleman Report [Coleman et. al. 1966], and has been true every year thereafter.

<sup>2</sup>This could be due, in part, to spatial mis-match [Kain 1968].

<sup>3</sup>There is no consensus on this view, however. A discussion follows in section II.

<sup>4</sup>See Fryer [2004] for a detailed discussion of these groups.

the ongoing puzzle of black underachievement.<sup>5</sup> The key idea is that individuals face a two audience signaling quandary: behaviors that promote labor market success are behaviors that induce peer rejection. The model involves an environment with three sorts of agents: individuals, firms, and a (suitably anthropomorphized) peer group. Individuals are endowed with a two-dimensional type: social and economic. The social type represents their value to the group and the economic type their value in the economic market. Individuals are also endowed with a unit of time to be allocated between schooling and leisure. After observing their two-dimensional type, individuals choose a fraction of time to devote to education which serves as a signal to firms about future productivity and to peers about social compatibility. Upon observing individual investments in education, the group decides whether or not to accept the individual and wages are set by firms who engage in Bertrand bidding to produce a homogeneous product.<sup>6</sup>

It is important to emphasize at the outset that, in the model, firms are assumed to have no interest in any employee's group membership, and groups are assumed not to have any basic preference over whether a potential member works hard at school, is employed or wealthy.<sup>7</sup> Consequently, there is no intrinsic conflict built into the model between individuals being highly educated and employed, and being members of a group; to the extent that such conflict emerges, therefore, it is as an equilibrium consequence of two-audience signaling. At the same time, other things equal, all social types strictly prefer to be accepted rather than rejected by their peer group. And, just as group acceptance is valuable to the individual, individuals yield value to the group through consumption externalities, community policing, so on and so forth. Peer groups, however, only

---

<sup>5</sup>There are many of other explanations put forth to explain racial differences in achievement, which range from the genetic inferiority of minorities [Jensen 1998 and Rushton 2000] to differences in neighborhoods and environment [Brooks-Gunn and Duncan, 1997]. All of these explanations fail on some dimension. The genetics argument runs in direct conflict with the substantial amount of biological evidence that fails to pin any difference in the biological taxonomy of human beings to race [Olson 2002 and Sykes 2002]. Environmental explanations do not account for the empirical observation that blacks still do substantially worse than whites in middle class suburban neighborhoods, where presumably the social and financial constructs are very similar [Ferguson 2001 and 2002].

<sup>6</sup>Several features of this framework are quite consistent with the motivation of recent work on optimal parochialism [Bowles and Gintis 2004] and identity [Akerlof and Kranton 2000]. Inter alia, these authors survey a litany of works, both academic and autobiographical, testifying to the tension many individuals of a minority culture feel between doing what is expected to remain accepted by their peers or social group (be it predicated on race, ethnicity, gender, or other network affiliation), and doing what is expected to succeed in a world dominated by those in the majority culture.

<sup>7</sup>This is an important point of departure from the standard explanations in the sociology and anthropology literatures.

want to accept members who are socially compatible group members in that they can be depended upon to support the group in difficult times. Examples are not hard to find; they range from gang members who can be trusted not to betray other members when subjected to police investigation, to residents of a community who can be relied upon to invest the time and effort to help their neighbors when they are in need [see Anderson 1990 and 1999, for more detailed examples].<sup>8</sup>

We establish three central results for equilibrium behavior satisfying a standard belief-based refinement. First, when either element of an individual's type (social or economic) is common knowledge, the equilibria of the model involve complete separation; all ability types choose distinctive levels of educational investment and only those with high social value are accepted. Second, when both elements of an individual's type are private information, no equilibria exist in which all types adopt distinct education choices; all equilibria must involve some pooling. Third and most interesting, equilibria involve a partition of individual economic abilities into at most three intervals: a (possibly empty) set of the lowest ability types and the set of highest ability types reveal themselves through a separating education strategy; ability types in the middle interval pool on a common education level. Only types in the lower intervals are accepted by the group. It is worth emphasizing that nothing is built into the model that requires accepted types to adopt a common educational investment; it is an equilibrium outcome. This partition produces novel predictions as one varies the wage structure, group size and value of membership, and the types of social interactions involved. The ability of the boundary types are strictly increasing in the value of group membership, which is likely higher in places with more racial segregation. Similarly, as the opportunity cost of group membership increases, the marginal ability types strictly decrease and the pooling level of education among the accepted types decreases. Ironically then, in environments in which 'acting white' is salient, improved external labor markets have the effect of encouraging more individuals to leave the group, while causing those in the group to invest less in education.

---

<sup>8</sup>An important characteristic of these and many other examples, one that in large part defines what it is to be a member of social group rather than a strictly economic market, is that the costs of membership are in terms of personal time and effort, not money per se. Although the assumption that all individuals prefer to be accepted by their peers is taken as primitive (and predicated on the sociological and psychological evidence that such preferences exist and are widespread [Freedman, Carlsmith, and Sears 1970]), the operationalization of which combinations of social and economic types constitute acceptable group members is endogenous to the model, giving rise naturally to a notion of peer pressure.

## II. CONFLICTING EVIDENCE ON ‘ACTING WHITE’

The notion that blacks view a set of distinctive behaviors (upward mobility, particular speech patterns, acquiescence to white authority) as “selling out,” “acting white,” or “Tomism” can be traced (at least) to the well-documented strife between house and field Negroes in everyday plantation life, and seems to have taken hold by the advent of the Black Power movement in the late 1960’s.<sup>9</sup> For instance, Van Deburg [1992] reports popular black literary artists depicting “Toms” having their tongues cut out for talking like white people. More recently, there has been a renewed interest in the idea as a plausible explanation of black-white differences in educational achievement. This is due, in part, to Fordham and Ogbu [1986] and Fordham [1996] who argue for the prevalence of an oppositional culture among black youth that eschew behaviors traditionally seen as the prerogative for whites.<sup>10</sup> Their hypothesis states that the observed disparity between blacks and whites stems from the following factors: (1) white people provide them with inferior schooling and treat them differently in school; (2) by imposing a job ceiling, white people fail to reward them adequately for their academic achievement in adult life; and (3) black Americans develop coping devices which, in turn, further limit their striving for academic success.<sup>11</sup> This dilemma between racial authenticity and achievement has been documented in many ethnographies and the popular media, though there is no consensus.<sup>12</sup>

There is an apparent conflict between the ethnographic evidence on ‘acting white’ and nationally

---

<sup>9</sup>A (suitably modified) version of ‘acting white’ is also prevalent in ethnographies involving the Buraku Outcasts of Japan [Devos and Wagasutma 1966], Italian immigrants in Boston’s West End [Gans 1962], the Maori of New Zealand [Chapple, Jefferies, and Walker 1997], Blacks on Chicago’s south side circa 1930 [Drake and Cayton 1945], the working class in Britain [Willis 1977], among others. In all cases high achievers receive a derogatory label from their peer group. For example, in the peer group society documented in Gans [1962], upward mobile youth interested in education were labeled “mobiles” and “sissies.” See Fryer [2004] for a detailed discussion of these groups.

<sup>10</sup>Generally, there are large literatures concerning group influences on individual decision-making in sociology and social psychology, yet efforts to develop more formal models addressing how such influences affect economic decisions in general, let alone with regard to education and investment in human capital, are relatively new. And within the formal literature, most of the work is devoted to understanding the economic implications of (more or less) given social norms: recent examples include Akerlof [1976, 1980], Bernheim [1994], Lindbeck et. al. [1999], and Cole et. al. [1992]. While much of this literature bears in some way on the issue here, none of it directly considers the role of peer pressure on human capital formation.

<sup>11</sup>Fordham and Ogbu suggest the problem arose partly because white Americans traditionally refused to acknowledge that black Americans were capable of intellectual achievement, and partly because black Americans subsequently began to doubt their own intellectual ability, began to define academic success as white people’s prerogative, and began to discourage their peers, perhaps unconsciously, from emulating white people in striving for academic success.

<sup>12</sup>For recent work on the prevalence of ‘acting white’ among blacks, see Chin and Phillips [2004], Corwin [2001],

representative studies that find no justification for the oppositional culture hypothesis – attempting to dismiss ‘acting white’ as nothing more than an urban (or more precisely, ethnographic) legend. Cook and Ludwig [1998] is the leading example of a study that fails to find empirical justification for an oppositional culture among black youth [see also Ainworth-Darnell and Downey 1998].

There are essentially two conditions that must be satisfied for ‘acting white’ to be salient: peer group norms between blacks and whites have to be significantly different and these norms must influence the educational production function. Cook and Ludwig [1998] attempt to prove that the first condition fails by correlating self-reported measures of popularity with academic achievement among 10th graders. They find that black high achievers are no more unpopular than black low achievers and, therefore, acting white does not exist. The relevant question, however, is whether or not the popularity-achievement gradient is the same for blacks *and* whites.<sup>13</sup>

Using two nationally representative datasets (National Educational Longitudinal Study and National Longitudinal Survey of Adolescent Health), Fryer and Torelli [2004] show that there are large racial differences in the relationship between popularity and grades. Several stylized facts emerge from their analysis. First, ‘acting white’ exists and is remarkably robust across many specifications and subsets of the data.<sup>14</sup> That is, there exists a strong and statistically significant difference between blacks and whites in the relationship between popularity and grades. For high achievers, the magnitude of ‘acting white’ more than doubles. Second, ‘acting white’ is particularly salient in public schools and schools in which the percentage of black students is less than twenty, but non-existent among students from high socio-economic families or among those that attend private schools. In schools that are less than twenty percent Black the ‘acting white’ coefficient is nearly three times as large as that in the overall sample. This is not consistent with Fordham and Ogbu [1986], but is a clear prediction of the model to be put forth here. Third, Fryer and Torelli [2004] estimate that eliminating the difference between blacks and whites in the relationship between

---

Fordham [1991], Ogbu and Davis [2003], Suskind [1998], or Tyson, Darity, and Castellino [2004]. One can also conduct a Lexis-Nexis search of major newspapers, which will yield scads of articles. Fryer [2004] documents over 20 groups across the world that face the dilemma between racial/cultural authenticity and economic achievement.

<sup>13</sup>Extrapolating anything from the “level” of questions assessing self-damaging information in a survey is deeply problematic. It is a standard measurement error issue. The popularity measured in a survey is some function of an individual’s true popularity and noise. If we assume that the noise is distributed similarly across racial groups, there is information in the racial difference in popularity – even when the level has no content.

<sup>14</sup>The Add Health data are ideal for understanding ‘acting white,’ as they contain more objective measures of popularity – asking students actually to list their friends. Fryer and Torelli [2004] calculate measures of popularity for each student by adding up all individuals who listed that student as a friend.

popularity and grades would account for roughly 10% of the black-white test score gap, 40% of the effort gap, and 60% of the difference between Blacks and Whites in time spent on homework. While the results do not reconcile all the differing opinions in the sociology and cultural anthropology literatures on the prevalence and impact of ‘acting white’, it is clear that the relationship between popularity and achievement is quantitatively different among Blacks compared to Whites, and can explain a significant fraction of the Black-White achievement gap along several dimensions.<sup>15</sup>

### III. A TWO-AUDIENCE SIGNALING MODEL

In what follows, we present a simple and stylized model of peer pressure that abstracts from all but the bare essentials necessary to illustrate the motivating ideas.<sup>16</sup> The key innovation is that an individual’s educational investment is a signal, both to potential employers about the individual’s productivity and to peers about the individual’s social compatibility. Further, while employers are free to adjust wages continuously in an individual’s signaled productivity, the peer group simply makes a binary decision regarding whether the individual is deemed acceptable or not. So, although it is assumed that employers have no direct interest in the individual’s social status and that peers have no direct interests in the individual’s productivity, the equilibrium consequence of two-audience signaling with a common decision is that a subset of productive types underinvest in education relative to the situation without any peers to impress. The following model captures this intuition.

#### III.A. Agents and actions

Consider three sets of agents: individuals, firms, and a (suitably anthropomorphized) peer group. An individual’s type is a pair,  $\tau = (\gamma, \theta) \in \{h, l\} \times [\underline{\theta}, \infty)$  with  $\underline{\theta} > 1$  and  $h > l > 0$ . Each component of an individual’s type is private information to the individual and assumed to be independently drawn. Let  $p(h)$  be the probability that  $\gamma = h$  and assume that  $\theta$  is chosen according to a smooth cumulative distribution function (CDF)  $F$  with density having full support

---

<sup>15</sup>It is important to also emphasize that the effects of ‘acting white’ reach far beyond investments in education. For example, it is likely to be salient in the production of informal associations, friendships, network affiliations, and other social relationships individuals obtain. These variables are highly correlated with economic success [see Calvo-Armengol and Jackson 2004].

<sup>16</sup>In earlier versions of this paper [Austen-Smith and Fryer 2003] we solve a substantially more general infinite horizon model with endogenous peer pressure.

on  $[\underline{\theta}, \infty)$ . Both  $p$  and  $F$  are assumed common knowledge. The interpretation is that  $\gamma$  is the individual's *social type*, reflecting his compatibility to the group, and  $\theta$  is the individual's *economic type*, reflecting his or her intrinsic ability or market potential. With some abuse of terminology, we refer to an individual of type  $\tau$  simply as "individual  $\tau$ ".

Individuals are endowed with one unit of time to allocate between leisure and education. Since education and time expended on acquiring education are identified without loss of generality here, let  $s \in [0, 1]$  denote the level of education acquired by an individual. Education is costly. We assume the costs can be separated into direct and indirect components. The indirect component of cost is through leisure forgone and we discuss this below. The direct component of cost depends on the individual's ability,  $\theta$ , and reflects the usual intuition that more able individuals can absorb education more easily than those less able in this respect. Let  $c(s, \tau) \equiv c(s, \theta) \geq 0$  be the direct cost to type  $\tau = (\gamma, \theta)$  of acquiring education  $s$ . As in the canonical signaling literature [Spence 1974], assume the cost function  $c$  is strictly increasing and convex in education, strictly decreasing in economic type and satisfies the single-crossing property. In symbols:

$$(1) \quad c_s > 0, c_{ss} > 0, c_\theta < 0; \text{ and } c_{\theta s} < 0,$$

where subscripts denote partial derivatives and  $c(0, \cdot) = 0$  as usual. Further, it is convenient to assume that, for all  $\theta$ ,  $\lim_{s \downarrow 0} c_s(s, \theta) = 0$  and  $\lim_{s \uparrow 1} c_s(s, \theta) = \infty$ , so that all types choose interior education levels.

Unlike types, education levels  $s$  are commonly observed and competitive bidding between firms leads to post-school employment at an endogenously determined wage-rate,  $w \geq 0$ . Competitive firms engage in Bertrand bidding for employees to produce a homogenous product under constant returns. An employee is characterized by a pair  $(s, \theta)$  describing the individual's education and economic type. The marginal product of an employee  $(s, \theta)$  to any firm is  $w(s, \theta) \geq 0$ . Assume marginal products are strictly increasing in education and economic type, with a nonnegative interaction between the two arguments. In particular, let

$$(2) \quad w(s, \theta) = \theta s.$$

Firms have no interest in any individual beyond his or her capacity as an employee; they neither observe nor care about an individual's social type or whether he or she belongs to a group. Firms choose wage-offers to maximize expected profit. And, the wage paid in equilibrium to an employee,  $(s, \theta)$ , under complete information on  $\theta$  is  $w(s, \theta)$ . Under incomplete information on  $\theta$ , however, education can serve as a signal of economic type. The wage offered to any potential employee is

that individual’s expected marginal product conditional on his or her observed education level and the educational investment strategy.

We are not interested here in details of the equilibrium wage-offer schedule. As such, from this point forward, we simply assume that individuals are paid a wage equal (in equilibrium) to their expected marginal product as just described and, where appropriate, specify the firms’ responses to any out-of-equilibrium action by an individual.

To this point, the setup essentially describes a canonical Spence signaling model; we now introduce a variation through the social component of individual types. The key innovation is that an individual’s educational choice problem is influenced by his or her peer group. Other things equal, being an accepted member of the group is assumed to be valued by all types. Assume that the value of leisure to an accepted individual with education  $s$  is given by  $(1 + \gamma)(1 - s)$ , whereas the value of leisure if individual  $\tau$  is rejected by the group is simply  $(1 - s)$ . The payoff to  $\tau$  of choosing education  $s$  and receiving wage  $w$  is

$$(3) \quad u(w, \alpha, s, \tau) = (1 + \alpha\gamma)(1 - s) + w - c(s, \theta),$$

where  $\alpha = 1$  if and only if the individual is accepted by the group;  $\alpha = 0$  otherwise.<sup>17</sup>

The (unitary actor) peer group has no direct interest in an individual’s economic potential.<sup>18</sup> Rather, the group is interested in accepting individuals whose social value to the group is sufficiently high, where social type and economic type are uncorrelated in the model. This is a crucial difference between an economic approach to understanding ‘acting white’ and that used in sociology and cultural anthropology. In the latter, groups are assumed to have preferences which are oppositional toward economic success (Fordham and Ogbu [1986]).

Let the group’s value from an individual of social type  $\gamma$  be  $v(\gamma)$  and assume  $l$  is sufficiently low that  $v(l) = 0 < v(h) = 1$  for all  $s > 0$ . Assume further that the group’s payoff from rejecting any individual is fixed,  $\hat{v} \in (0, 1)$ .<sup>19</sup> The group is only interested in accepting “high” social types and has no direct concern with such types’ educational levels. On the other hand, an individual’s

---

<sup>17</sup>We assume linearity for convenience. All forthcoming results hold if the utility function is: twice differentially concave; increasing in leisure; both total and marginal values from consuming any strictly positive amount of leisure are greater as an accepted group member than otherwise; and zero leisure is worthless irrespective of group acceptance.

<sup>18</sup>Of course this is an extreme assumption. It is, however, the conservative assumption to make for our purposes.

<sup>19</sup>This value can be derived from a multi-period model of social interaction between the group and the individual: see, for instance, Fryer [2004]. Further, the assumption that there are only two social types is not essential. Our principal qualitative results go through with a continuum of social types and payoffs to the group,  $v(\gamma)$ , strictly increasing in  $\gamma$ .

choice of education in principle conveys information to the group regarding the individual's social type.

Putting the pieces together, individuals first choose their education according to a strategy  $\sigma : \{h, l\} \times [\underline{\theta}, \infty) \rightarrow [0, 1]$ , where  $\sigma(\tau) = s \geq 0$  denotes the level of education achieved by  $\tau$  under  $\sigma$ ; given  $\sigma$ , the firms make a wage-offer  $\omega(s) = sE[\theta|\sigma(\tau) = s] \geq 0$  to an individual with education level  $s$ ; and, letting  $\alpha(s)$  denote the probability that an individual with education level  $\sigma(\tau) = s$  is accepted by the group,  $\alpha(s) = 1$  if and only if  $\alpha(s) \geq \hat{v}$ . To insure the group has a nontrivial problem, assume the prior that any individual is a high social type,  $p(h)$ , is strictly less than  $\hat{v}$  and that the difference between the high and the low social types is non-negligible.<sup>20</sup> It is important to note that any racial or ethnic differences in the relationship between achievement and social acceptance comes through  $\hat{v}$ .

We are interested in how the desire to be accepted by one's peer group, albeit modeled in the minimal fashion described above, influences an individual's choice of education. The solution concept, much-used for analyzing costly signaling games, is the D1 refinement of sequential equilibrium [Banks and Sobel, 1987; Cho and Kreps, 1987; Cho and Sobel 1990], suitably extended to cover our two-audience setting. It is well-known that there is a unique D1 equilibrium in the canonical educational signaling game without a group; all types separate with education strictly increasing in type, and firms pay individuals their true marginal product [Riley 1979]. Moreover, to avoid lower types mimicking higher types, the separating strategy requires all but the lowest type to over-invest in education relative to their complete information utility-maximizing choice. In our model, however, an individual's type consists of two independent components, social and economic, and there are two distinct agents separately interested in these components. Extending the D1 refinement to our setting is straightforward. First, by definition of a sequential equilibrium, firms and the group all form out-of-equilibrium beliefs identically. Second, these uninformed agents form their beliefs consequent on an out-of-equilibrium message by identifying those types having the most incentive (in a set-inclusion sense) to offer the deviation, conditional on the uninformed agents' best-responding to these beliefs, where the identification of the types having most to gain from any deviation requires consideration of both components of an individual's type,  $\tau$ .<sup>21</sup>

---

<sup>20</sup>Exactly what this means is made precise in the next section.

<sup>21</sup>More precisely, suppose that, for all responses by the group and the firms, whenever some type  $\tau$  is weakly better off relative to his or her equilibrium payoff by choosing a particular out-of-equilibrium action  $s$ , then a type  $\tau'$  is strictly better off choosing  $s$ . Then D1 requires that both the group and the firms assign zero probability to the deviant being type  $\tau$  rather than type  $\tau'$ .

### III.B. ‘Acting white’ equilibrium

It is immediate from the assumptions on the utility of leisure and the cost function  $c$ , that no type ever chooses the maximum available education level,  $s = 1$ ; thus  $(1 - s) > 0$  in any equilibrium with or without complete information. In particular, if there is complete information on  $\tau$ , individuals are paid a wage  $w = w(s, \theta)$  and (3) implies that for all  $\tau$  and all  $s \in [0, 1]$ ,

$$(4) \quad u(w, 1, s, (\theta, \gamma)) > u(w, 0, s, (\theta, \gamma)).$$

Also,  $u(w, \alpha, s, \tau)$  is strictly concave in  $s$  with utility maximizing education level  $s_\alpha^c(\tau)$ . The first order condition defining an interior level of education  $s_\alpha^c(\tau)$  is

$$(5) \quad \frac{du}{ds} = [\theta - (1 + \alpha\gamma)] - c_s(s, \theta) = 0.$$

It is convenient (but not essential) to normalize types so that

$$(6) \quad \underline{\theta} = (1 + h).$$

Then, given complete information and  $\alpha \in \{0, 1\}$ , we have  $s_1^c(\underline{\theta}, h) = 0 < s_1^c(\underline{\theta}, l)$  and, for all  $\theta > \underline{\theta}$  and all  $\gamma$ ,

$$(7) \quad 0 < s_1^c(\theta, h) < s_1^c(\theta, l) < s_0^c(\theta, \gamma) \equiv s_0^c(\theta).$$

We suppose throughout that the difference between the high and the low social type is sufficiently large that an individual of type  $(\underline{\theta}, l)$  prefers to be rejected by the group with income  $w = \underline{\theta}s_0^c(\underline{\theta})$  to being accepted by the group with no income; specifically, assume

$$(8) \quad u(w(s_0^c(\underline{\theta}), \underline{\theta}), 0, s_0^c(\underline{\theta}), (\underline{\theta}, l)) \geq u(0, 1, 0, (\underline{\theta}, l))$$

or, equivalently,

$$s_0^c(\underline{\theta})h - c(s_0^c(\underline{\theta}), \underline{\theta}) \geq l.$$

It is useful to consider two benchmark cases. In the first, social types are common knowledge but economic types are private information. Then only high social types are accepted by the peer group in equilibrium and the unique D1 equilibrium involves all economic types separating according to the Riley strategy; that is,  $\sigma$  is strictly increasing in  $\theta$  with  $\sigma(\underline{\theta}, l) = s_0^c(\underline{\theta})$ ,  $\sigma(\underline{\theta}, h) = s_1^c(\underline{\theta}, h)$  and, for all  $\theta > \underline{\theta}$ ,  $\sigma(\theta, \gamma) > s_\alpha^c(\theta, \gamma)$ . Moreover, low social types invest strictly more in education than high types. See Figure I.

Figure I here

Note that while there can be two distinct economic types choosing a given level of education in the equilibrium, the fact that social type is common knowledge permits firms to separate these economic types and pay each their respective marginal products.

The second benchmark has economic types common knowledge and social types private information. In this case, there exists a D1 equilibrium separating in social types.

**Proposition I:** If economic types  $\theta$  are common knowledge but social types  $\gamma$  are private information, then there exists a D1 equilibrium separating in social type such that, for all economic types  $\theta$ ,  $\sigma(\theta, l) = s_0^c(\theta)$  and  $\sigma(\theta, h)$  is strictly increasing in  $\theta$  with  $\sigma(\theta, h) \in [0, s_1^c(\theta, h)]$ .

In both benchmark cases, therefore, complete information concerning one component of individual type suffices to admit separating equilibria in the other component and, further, individuals' educational investments need not be less than they would be under complete information on types  $\tau$  (although, depending on details of the cost function  $c$ , some underinvestment might be necessary to insure low social types do not mimic high social types). This is not so surprising. With only one component of type private information, individuals face a single-audience signaling problem. Things change when both components of type are private information, however, and an individual has to resolve a two-audience signaling problem. Hereafter, assume that both  $\gamma$  and  $\theta$  are private information.

Let  $\tilde{s}(\underline{\theta}) \in (0, s_1^c(\underline{\theta}, l))$  be the education level at which  $(\underline{\theta}, l)$  is indifferent between group rejection with  $s_0^c(\underline{\theta})$  and group acceptance at  $\tilde{s}(\underline{\theta})$ :

$$u(w(\tilde{s}(\underline{\theta}), \underline{\theta}), 1, \tilde{s}(\underline{\theta}), (\underline{\theta}, l)) = u(w(s_0^c(\underline{\theta}), \underline{\theta}), 0, s_0^c(\underline{\theta}), (\underline{\theta}, l)).$$

That  $\tilde{s}(\underline{\theta})$  exists follows from (8), concavity and continuity of (full information) payoffs in education. Say that an education strategy  $\sigma$  is *separating in  $\theta$*  on an interval  $I$  if, for all  $\gamma$  and all distinct economic types  $\theta, \theta'$  in  $I$ ,  $\sigma(\theta, \gamma) \neq \sigma(\theta', \gamma)$ . The following proposition states our main result.

**Proposition II:** D1 equilibria exist and, if  $\sigma(\tau)$  is a D1 equilibrium strategy, it satisfies the following properties.

(I) For all types  $\tau = (\theta, l)$ ,  $\sigma(\tau)$  is the unique efficient strategy separating in  $\theta$  on  $[\underline{\theta}, \infty)$  with initial condition  $\sigma(\underline{\theta}, l) = s_0^c(\underline{\theta})$ .

(II) There exist economic types  $\theta_L \geq \underline{\theta}$  and  $\theta_H > \theta_L$  such that, for all types  $\tau = (\theta, h)$ ,

- (a) if  $\theta \in [\underline{\theta}, \theta_L]$ , then  $\sigma(\tau)$  is the unique efficient strategy separating in  $\theta$  on  $[\underline{\theta}, \theta_L]$  with initial condition  $\sigma(\underline{\theta}, h) = 0$ ;
  - (b) if  $\theta \in (\theta_L, \theta_H]$ , then  $\sigma(\tau) = \hat{s}$  where  $\hat{s} \in [\sigma(\theta_L, h), \tilde{s}(\underline{\theta})]$ ; and
  - (c) if  $\theta > \theta_H$ , then  $\sigma(\theta, h) = \sigma(\theta, l)$ .
- (III) The group accepts an individual if and only if  $s \leq \hat{s}$ .

Figures II illustrate the possible equilibria.

Figures II here

Proposition II admits the possibility that the lowest accepted types might adopt distinct education levels in equilibrium. However, the highest accepted economic types necessarily pool on a common level of education. And since the highest possible such level of education chosen by any accepted type lies strictly below the level that the least able low type,  $(\underline{\theta}, l)$ , would choose if accepted under complete information, this highest possible level of education is necessarily very low indeed. We expect the variance in education among accepted group members to be negligible. It is worth emphasizing that, although the group has no intrinsic interest in any individual's economic ability and although social and economic types are independent, it is only the relatively low economic types  $(\theta, h)$  who select into the group by underinvesting in their education. The equilibrium connection between the social and economic types lies in the opportunity cost of underinvesting in education. For sufficiently high economic types, the income forgone by insuring group acceptance is simply too great.

Comparing Proposition II to the two benchmark cases, it is clear that the conflicting incentives induced by signaling to two audiences leads to underinvestment in education among low and moderate economic types. Low social types are always rejected by the group and invest appropriately (that is, they overinvest to separate themselves in economic type); high social types, however, can underinvest significantly to insure group acceptance. For sufficiently able economic types, the opportunity cost of group acceptance becomes excessive and they choose to signal economic ability to firms rather than social type to peers. It is worth noting too, that given economic types can consistently adopt distinct education levels in equilibrium, depending on whether they are accepted by the group.

### III.C. Comparative statics

Perturb the specification of payoffs slightly by letting an individual's productivity be  $\beta w(\theta, s)$  where  $\beta > 0$ . The parameter  $\beta$  can be interpreted either in terms of a neutral change in worker productivity (with  $\beta \geq 1$ ) or as a discount factor (with  $\beta > 0$ ). In the latter case, we imagine educational investment and peer group activity occurs during the 'school years' but wages are earned only during the 'post-school years'. Then it is straightforward to confirm the following comparative statics (stated without proof).

**Proposition III:** Let  $\sigma$  be a D1 equilibrium strategy satisfying the properties described in Proposition 2 with  $\hat{s} = 0$ . Then, other things equal:

- (I) a marginal increase in  $\beta$  induces a fall in  $\theta_H$  and an increase in the educational achievement and wages of all rejected types;
- (II) a marginal increase in  $h$  induces an increase in  $\theta_H$ .

Interpreting  $\beta$  as a discount factor, the first observation states that the less individuals value the present over the future, the more attenuated peer pressure becomes. Similarly, if  $\beta$  measures a shift in worker productivity, the observation is that an increase in the economic opportunity cost of group membership results in fewer types being willing to sacrifice education for peer acceptance. On the other hand, the second observation states that the greater the high social type values group acceptance relative to income, the more effective is the desire to impress the group of one's social acceptability. These implications are intuitive: as the value of group acceptance increases, the ability of the boundary types ( $\theta_L$  and  $\theta_H$ ) increase. When the maximal education level among accepted group members is not zero, there might be several triples  $(\theta_H, \theta_L, \hat{s})$  with  $\hat{s} > 0$  consistent with equilibrium behavior. In this event, the comparative statics above continue to obtain for the equilibrium with a fixed  $\hat{s}$ . In particular, both boundary points, ( $\theta_L$  and  $\theta_H$ ) fall with an increase in  $\beta$ , as does the average wage of those accepted by the group relative to those rejected by the group.<sup>22</sup>

Empirically, the value of group acceptance (as reflected in  $\gamma$ ) is a function of many social variables including segregation, crime, neighborhood structure and density, and so on. In the next

---

<sup>22</sup>Intuitively, given a fixed  $\hat{s} > 0$ , the first-order impact of a marginal increase in  $\beta$  is to shift relative wages so that  $\theta_H$  strictly prefers to separate and be rejected rather than pool on  $\hat{s}$  and be accepted, while  $\theta_L$  strictly prefers to pool on  $\hat{s}$  rather than separate within the group. Thus both boundary types shift downward, inducing a second-order reduction in the wage level of those choosing  $\hat{s}$  and resulting in a fall in the average wage of those accepted by the group to those rejected.

section, we consider the implications of our results for some of the more significant influences on the worth of group acceptance. Of necessity this will be more informal, as we have not explicitly derived the value of  $\gamma$  from these underlying characteristics. Before doing this, however, it is useful to note that although we have stated the comparative statics in terms of  $\beta$  and  $h$ , similar implications apply with respect to analogous changes in the opportunity costs of group membership (for example, a change in the direct cost of acquiring education). In general, an increase in the opportunity costs of group membership has two effects: an increase in the set of rejected types and a (weak) decrease in time or effort allocated to education by members of the accepted group. This latter comparative static illustrates a distinct bipolarization in equilibrium behavior as a result of changes in the economic environment.<sup>23</sup> If the costs of being a group member increase, the marginal type eschews the group and adopts correspondingly higher education and wages, whereas those who adhere to group norms become even more marginalized. There is some suggestive evidence that this type of bipolarization occurred among blacks in the 1980s when the price of skilled labor increased substantially. Rubinstein [2001] documents that while the variance in the wage distribution among blacks was relatively constant from 1960-1980 – there was a substantial increase in the 1980’s with high income blacks converging to whites and low income black diverging. More generally, costs of group membership might increase for a myriad of reasons including: increases in gang violence; decreases in segregation; or an increase in the returns to skill in the labor market. And, as this cost tends towards infinity, peer pressure ceases to exist. Indeed, this seems to be precisely what happened, in effect, to the Italian immigrant community in Boston’s West End. The gentrification of the West End circa 1958 increased the cost of remaining in, what Gans [1962] refers to as, “the peer group society” substantially. As a result, the accusations of ‘acting mobile’ became less salient and assimilating behaviors increased.

#### IV. INTERPRETING BLACK UNDERACHIEVEMENT THROUGH THE LENS OF ‘ACTING WHITE’

In this section, we attempt to understand some of the patterns and paradoxes of black underachievement through the lens of the model put forth in the preceding section. The evidence is not meant *prove* whether or not ‘acting white’ is an important economic phenomenon, as most of the theoretical implications are consistent with other models of self selection. More modestly, we endeavor to show that the model’s predictions are consistent with empirical facts on the decline of

---

<sup>23</sup>Fryer [2004] produces a similar bipolarization in a infinitely repeated game setup.

black neighborhoods in the 1970's, black underachievement in middle class neighborhoods, and the efficacy of particular educational and job interventions.

#### **IV.A The Rise of the Ghetto**

From American Chattel slavery through Jim Crow, the value of racial solidarity and group membership was extremely high. In a typical black community, doctors, lawyers, postmen, and others with lower occupational status, lived in the same vicinity. High educated blacks were just as likely as low educated blacks to fulfil their obligations to the community. With the decrease in institutional discrimination and the increase in housing integration came many new opportunities, including the choice of opting out of the group. With this change, community monitoring of agents' educational investments became an important predictor of their group loyalty. In his classic portrayal of neighborhoods in Chicago, Wilson [1978] argues that the African American community was splitting into two, with middle class blacks increasing their position relative to whites, and poor blacks becoming even more marginalized. Wilson's conjecture is that the plight of the black inner cities was due to the erosion of their social networks and social capital. This element of self-selection is readily seen in the framework presented here, as we increase the cost of group membership. There are, however, some subtle differences. Whereas Wilson argues that networks are to blame, we argue that the very presence of high ability blacks and institutional barriers gave incentives to those on the margin to invest in group membership. This observation is consistent with the findings involving 'acting white' in suburban schools [Fryer and Torelli 2004]. Essentially, Proposition II shows that as one increases the porosity between black and white social interactions, negative peer sanctions become more salient. Thus, our hypothesis for the erosion of inner-cities is slightly different from Wilson [1978]. Wilson argues that the corrosion of social networks and role models are to blame. We assert that the very presence of institutional discrimination and the lack of mobility eliminated the two-audience signaling problem and, hence, the link between education and 'acting white.'

#### **IV.B. The Middle Class Paradox**

While the acting white hypothesis may explain sub par academic performance in low-income black neighborhoods, one potential puzzle is the black middle-class. It is well documented that black adolescents in middle class neighborhoods are not achieving academically at the same rate as their white counter parts [Pattillo-McCoy 1999, Ferguson 2002]. And, acting white is more prevalent in suburban schools than predominantly black schools [Fryer and Torelli 2004, Ogbu and Davis

2003]. This has puzzled many, since presumably black and white children in these neighborhoods have been reared under similar conditions. To understand this paradox, one has to consider the impact of racial segregation in housing. Due to a peculiar history, the black middle class are much more likely to live in neighborhoods that border poor black neighborhoods.<sup>24</sup> Jargowsky and Bane [1991] show that black middle class neighborhoods are much more likely to create a buffer zone between the black poor and white non-poor. Massey and Denton [1993] report that blacks with college educations have more than a 20 percent chance of coming in contact in their neighborhood with someone receiving welfare, whereas college-educated whites have an 8 percent chance. This pattern was repeated for interaction with blue-collar workers, high school drop-outs, and the unemployed. Because the nature of social interactions differs substantially between blacks and whites with similar incomes, the relative “group” is quite different. This makes racial loyalty of middle class blacks meaningful – which could explain lower achievement of middle class black students relative to their white peers. An important implication of this being that if black high achievers could limit their interactions to only other high achievers, ‘acting white’ would no longer be salient. That is, there is a non-monotonic relationship between percentage of black interactions and the saliency of peer group norms. At the extremes (high or low) negative peer sanctions are minimal. In the intermediate range, where behaviors serve as credible signals of loyalty, ‘acting white’ is most meaningful.

#### **IV.C. Policy Interventions**

There have many attempts at closing the black-white achievement gap. Some of these interventions involve taking students out of their classroom environments and enrolling them in programs away from their communities. This usually takes one of two forms: students are removed from their community for a fixed amount of time (day, summer, etc.), or students are removed from their community permanently. Versions of the latter characterization include the Gautreaux program, a major initiative adopted by the courts to provide a metropolitan-wide remedy for racial discrimination in Chicago’s public housing, Moving-to-Opportunity, Job Corps, and the A Better Chance Program.<sup>25</sup> In the A Better Chance initiative, students leave their families and live with a host family to attend better schools. Consider the following quote from a student in the A Better Chance Program.

---

<sup>24</sup>See Pattillo-McCoy [1999], chapter 2, for a nice discussion on the evolution of the black middle class.

<sup>25</sup>To read more about this program, visit <http://www.abetterchance.org>

“I felt I could be more involved with my studies here [in the host family]. At home, I would be distracted by peer pressures to hang out, smoke and drink. Here, I can focus on the academics. You face peer pressure wherever you go, but at Radnor there are more kids into their studies.<sup>26</sup>”

Our two audience signaling model predicts that these types of interventions will have larger marginal effects on students’ educational achievement because they change the nature of group interactions. Other interventions induce sorting, but have an ambiguous effect on the marginal student. The available evidence suggests that this is indeed the case. Sixty-five percent of the students in the A Better Chance program come from single parent families and thirty-three percent of them are beneath the poverty threshold; however, ninety-nine percent of the A Better Chance seniors immediately enroll in college. This is significantly larger than any other secondary educational intervention.<sup>27</sup>

More convincing are the results from analysis of the Job Start and Job Corps programs. Job Corps is the nation’s largest and most comprehensive residential, education and job training program for at-risk youth, ages 16 through 24. It takes the students to (predominantly rural) training centers where they receive free room and board along with intense training in one of 100 vocational specializations. Conversely, Job Start uses the same teaching curricula as Job Corp, but the students stay at home and commute to a local training site. As the model would predict, Job Corps has larger effects. It has been shown that Job Corp increases earnings and reduces crime, whereas, Job Start has shown statistically insignificant effects.<sup>28</sup>

If ‘acting white’ is suppressing achievement, the aim of any intervention must be geared either toward changing the nature/definition of group interactions or eliminating them completely – which have different implications for policy depending on which is pursued.<sup>29</sup> A nice test of the efficacy of the latter policy would be to investigate the black-white achievement gap among military children

---

<sup>26</sup> “Going ‘Away’ to School in Radnor: A Better Chance for Teens Who Put Their Schooling First.” The Suburban and Wayne Times, September 23, 1999.

<sup>27</sup> There are thorny selection issues to consider before one can truly test the programs causal effects.

<sup>28</sup> See <http://www.jobcorps.org> and [http://www.mdrc.org/project\\_9\\_60.html](http://www.mdrc.org/project_9_60.html) for results on Job Corp and Job Start programs, respectively.

<sup>29</sup> This could partially explain the anomalous results from the Moving-to-Opportunity Experiments where it was found that moving poor families to better neighborhoods had no effect on achievement for girls and a modest negative effect for boys (Kling, and Liebman 2004). And, boys were more likely to be in trouble with police. If ‘acting white’ was salient, then these results are consistent with a model in which the youth found it necessary to signal to their old peer group that they were worthy of acceptance.

at military schools, where presumably, racial differences in the relationship between popularity and grades are minimal. The theory predicts that these gaps will be substantially smaller. As evidence to this effect, Brown [2000] and Smrekar et. al [2001] show that while white children from military families score slightly higher than do their civilian counterparts, black children from military families do significantly better. The black-white test score gap is about 40 percent smaller in the military than in civilian schools. The Hispanic-white gap is approximately 60 percent smaller.

## V. CONCLUSION

Nearly 140 years after the abolishment of slavery and fifty years after the end of legalized disenfranchisement, many economic indices exhibit drastic racial inequities. Particularly stark are differences in educational achievement. We study one possible explanation for some of these disparities, ‘acting white’. The model here formalizes a conception of ‘acting white’ and explores its implications for individuals’ education decisions during their school years. Together, two of the main results from the model yield the motivating stylized facts regarding ‘acting white’ and underachievement documented in the anthropology and sociology literatures regarding low income populations across the world: first, there exist no equilibria in which all types of individuals adopt distinct educational investment levels and, second, all equilibria satisfying a standard refinement involve a partition of the type space in which all types accepted by the group pool on a common low education level and all types rejected by the group separate at distinctly higher levels of education with correspondingly higher wages.

Two further points are worth emphasizing. First, ‘acting white’ is not unique to Blacks and we (purposefully) reference other groups plagued by similar phenomena. Second, and most important, because various insidious forms of social interaction such as ‘acting white’ exist does not imply that nothing can be done about them. The comparative static results (Proposition III) suggest that improved labor markets, group incentives, and means for supporting implicit community-specific contracts are likely to undermine acting white.

In the past, the sorts of interaction explored here have been used by some to argue that particular subgroups are responsible for their own marginalization [see, for example, McWhorter 2000]. The implication being that policies aimed at eradicating educational achievement differences are ill-advised. In contrast, by analyzing such a phenomenon in a rational choice framework, it is transparent that the behavior is a result of strategic interaction for which any group with the same initial conditions would fall victim. As such, nothing should be ascribed to the inherent values, prefer-

ences, or ideologies of particular groups who are plagued by this insidious form of social interaction. Acting white is an equilibrium phenomenon; the consequence of two-audience signaling.

## APPENDIX: PROOFS OF PROPOSITIONS

This Appendix contains the formal arguments for the results in the text. The first two results confirm some standard properties of equilibria to signaling games.

**Lemma I:** Fix an equilibrium and let  $\tau, \tau'$  be any two types. Then:

- (I)  $\gamma = \gamma'$  and  $\theta > \theta'$  imply  $\sigma(\tau) \geq \sigma(\tau')$ ;
- (II) If  $\theta = \theta'$  and  $\gamma > \gamma'$  then either  $\alpha(\sigma(\tau')) = 0$  or  $\alpha(\sigma(\tau)) = \alpha(\sigma(\tau')) = 1$  and  $\sigma(\tau) \leq \sigma(\tau')$ .

**Proof.** In equilibrium, no type  $\tau$  finds it strictly better to mimic any other type. Thus, for all pairs  $\tau, \tau'$ ,

$$\begin{aligned} u(\omega(s), \alpha(s), s, \tau) &\geq u(\omega(s'), \alpha(s'), s', \tau) \text{ and} \\ u(\omega(s'), \alpha(s'), s', \tau') &\geq u(\omega(s), \alpha(s), s, \tau'), \end{aligned}$$

where  $s = \sigma(\tau), s' = \sigma(\tau')$  and so forth. Substituting for  $u(\cdot)$  from (3), we have,

$$\begin{aligned} (1 + \alpha\gamma)(1 - s) + w - c(s, \theta) &\geq (1 + \alpha'\gamma)(1 - s') + w' - c(s', \theta) \\ (1 + \alpha'\gamma')(1 - s') + w' - c(s', \theta') &\geq (1 + \alpha\gamma')(1 - s) + w - c(s, \theta') \end{aligned}$$

Adding the inequalities and collecting terms gives,

$$[c(s', \theta) - c(s, \theta)] - [c(s', \theta') - c(s, \theta')] \geq [\gamma - \gamma'][\alpha'(1 - s') - \alpha(1 - s)].$$

Assume  $\theta > \theta'$ . Then setting  $\gamma = \gamma'$  in this inequality and recalling  $c_{s\theta} < 0$  yields the first claim; and setting  $\theta = \theta'$  and  $\gamma > \gamma'$  yields the second.  $\square$

**Lemma II:** Consider any equilibrium and interval of types  $T$  such that  $\tau, \tau' \in T$  implies  $\gamma = \gamma'$  and  $\alpha(\sigma(\tau)) = \alpha(\sigma(\tau'))$ . Then  $\sigma$  separating in  $\theta$  on  $T$  implies  $\sigma(\tau) > s_\alpha^c(\tau)$  for all  $\tau$  interior to  $T$  and  $\alpha \in \{0, 1\}$ .

**Proof.** By hypothesis,  $\alpha$  is constant and  $\sigma(\tau)$  is separating in  $\theta$  for fixed  $\gamma$ . Hence, by Lemma I(I),  $\sigma(\tau)$  is strictly monotonic increasing in  $\theta$  on the relevant interval and, therefore, differentiable almost everywhere. Consequently, local incentive compatibility implies

$$\left. \frac{du(\omega(\sigma(\theta'), \gamma), \alpha, \sigma(\theta'), \gamma), \tau)}{d\theta'} \right|_{\theta'=\theta} = 0.$$

Since  $\sigma$  is separating on  $T$ ,  $\omega(\sigma(\theta), \gamma) = w(s, \theta)$  where  $s = \sigma(\theta, \gamma)$ . Substituting and doing the calculus yields,

$$\left[ (\theta - (1 + \alpha\gamma)) \frac{d\sigma(\tau)}{d\theta'} + \sigma(\tau) - c_s \frac{d\sigma(\tau)}{d\theta'} \right]_{\theta'=\theta} = 0$$

or

$$(9) \quad \left. \frac{d\sigma(\tau)}{d\theta'} \right|_{\theta'=\theta} = \frac{\sigma(\tau)}{c_s(\sigma(\tau), \theta) - [\theta - (1 + \alpha\gamma)]} > 0$$

where the inequality follows from Lemma I. Under complete information, however,  $s_\alpha^c(\tau)$  solves the first order condition, (5). Therefore, under complete information, the denominator of  $d\sigma/d\theta$  is zero. But since this derivative is strictly positive under incomplete information and separating, it must be that  $\sigma(\tau) > s_\alpha^c(\tau)$ .  $\square$

For  $\tau = (\theta, \gamma)$ , let  $u^*(\alpha, \tau) \equiv u(w(s_\alpha^c(\tau), \theta), \alpha, s_\alpha^c(\tau), \tau)$  denote  $\tau$ 's maximal utility under complete information conditional on  $\alpha$ .

**Proof of Proposition I.** Because  $\theta$  is common knowledge, in any equilibrium, the firm simply offers wage  $w(s, \theta)$  to any individual with education  $s$  and economic type  $\theta$ . For all economic type  $\theta$ , let  $\sigma(\theta, l) = s_0^c(\theta)$ . By (8),  $\tau = (\underline{\theta}, l)$  weakly prefers to be rejected with  $s_0^c(\underline{\theta}) > 0$  than accepted with  $s_1^c(\underline{\theta}, h) = 0$ . And by the assumptions on payoffs,  $u^*(\alpha, \tau)$  and  $s_\alpha^c(\tau)$  are strictly increasing in  $\theta$  with  $u(0, 1, 0, (\theta, \gamma)) = 1 + \gamma$ , all  $\tau = (\theta, \gamma)$ . It follows from the concavity and continuity of payoffs and (8), therefore, that for each economic type  $\theta$  there exists an education level  $\tilde{s}(\theta) \in (0, s_1^c(\theta, l))$  at which  $(\theta, l)$  is indifferent between group rejection with  $s_0^c(\theta)$  to acceptance with  $\tilde{s}(\theta)$ :

$$(10) \quad u(w(\tilde{s}(\theta), \theta), 1, \tilde{s}(\theta), (\theta, l)) = u(w(s_0^c(\theta), \theta), 0, s_0^c(\theta), (\theta, l)).$$

See Figure III.

Figure III here

Substituting from (3) and collecting terms, (10) can be written

$$(\theta - 1)(s_0^c(\theta) - \tilde{s}(\theta)) - [c(s_0^c(\theta), \theta) - c(\tilde{s}(\theta), \theta)] - l(1 - \tilde{s}(\theta)) = 0.$$

Differentiating through, collecting terms and using the first order condition (5) defining  $s_0^c(\theta)$ , yields

$$\begin{aligned} & \frac{d\tilde{s}(\theta)}{d\theta} [\theta - (1+l) - c_s(\tilde{s}(\theta), \theta)] \\ = & [s_0^c(\theta) - \tilde{s}(\theta)] - [c_\theta(s_0^c(\theta), \theta) - c_\theta(\tilde{s}(\theta), \theta)]. \end{aligned}$$

By (7),  $s_0^c(\theta) > \tilde{s}(\theta)$  and, by  $c_{s\theta} < 0$ ,  $c_\theta(s_0^c(\theta), \theta) < c_\theta(\tilde{s}(\theta), \theta)$ ; hence, the right side of the equality is strictly positive. And since  $\tilde{s}(\theta) < s_1^c(\theta, l)$ , (5) implies  $[\theta - (1+l)] > c_s(\tilde{s}(\theta), \theta)$ . Therefore,  $d\tilde{s}(\theta)/d\theta > 0$ .

For all types  $(\theta, h)$ , define  $\sigma(\theta, h) = \min\{\tilde{s}(\theta), s_1^c(\theta, h)\}$ . By concavity of  $u$ , all types  $(\theta, h)$  strictly prefer to be accepted with  $\tilde{s}(\theta)$  than rejected with  $s_0^c(\theta)$  and, by construction,  $(\theta, l)$  is indifferent between being rejected with  $s_0^c(\theta)$  than accepted with  $\tilde{s}(\theta)$ . Moreover, since  $d\tilde{s}(\theta)/d\theta > 0$ ,  $\sigma(\theta, h)$  is strictly increasing in  $\theta$ .

It remains to check that, under D1, there is no profitable deviation by any type  $(\theta, \gamma)$  to an education level  $s \notin \{\sigma(\theta, h), s_0^c(\theta)\}$ . In particular, it suffices to check only that there is no profitable deviation to some  $s \in (\sigma(\theta, h), s_0^c(\theta))$ . Suppose  $\sigma(\theta, h) = s_1^c(\theta, h)$  and consider a deviation to some  $s' \in (\tilde{s}(\theta), s_0^c(\theta))$ . Then  $\sigma(\theta, h) = s_1^c(\theta, h)$  implies that  $(\theta, h)$  is strictly worse off for all beliefs of the group but, by definition of  $\tilde{s}(\theta)$ ,  $(\theta, l)$  strictly profits from the deviation in the case that the group believes the deviant is a high social type and accepts the individual. Thus D1 requires the group believe the deviant is a low social type and rejects the individual, in which case  $(\theta, l)$  strictly prefers  $s_0^c(\theta)$  with the higher wage to the deviation  $s'$ . Now consider a deviation to some  $s' \in (s_1^c(\theta, h), \tilde{s}(\theta))$ . By definition of  $\tilde{s}(\theta)$  and  $s_1^c(\theta, h)$ , respectively, both the low and high social types are strictly worse off making the deviation to  $s'$  for all group beliefs. Therefore, no type has a profitable deviation in this case.

Suppose  $\sigma(\theta, h) = \tilde{s}(\theta)$ . If an individual deviates to some  $s' \in (s_1^c(\theta, h), s_0^c(\theta))$  then a preceding argument applies and the group surely rejects the deviant. So consider a deviation to  $s' \in (\tilde{s}(\theta), s_1^c(\theta, h)]$ . Then both types  $\gamma$  are strictly better off conditional on group acceptance and strictly worse off conditional on rejection. However, if the group believes either types is equally likely to have made the deviation and thus randomizes over whether to accept or reject the individual,  $\alpha(s') \in (0, 1)$ , the low type is willing to choose  $s'$  for a lower value of  $\alpha(s')$ . Hence, D1 requires the group to believe the low type deviates to  $s'$  and therefore rejects the deviant. Thus no type has any incentive to deviate in this case, completing the proof.  $\square$

Hereafter, assume both social type  $\gamma$  and economic type  $\theta$  are private information. Proof of Proposition II, our main result, rests on several lemmata.

**Lemma III:** If  $\sigma$  is a D1 equilibrium strategy then  $\sigma$  is separating in  $\gamma$  on some interval  $[\underline{\theta}, \hat{\theta})$  and  $\sigma(\underline{\theta}, h) = s_1^c(\underline{\theta}, h) = 0$ .

**Proof.** Suppose  $\sigma(\tau)$  is pooling in  $\gamma$  for all  $\tau = (\theta, \gamma)$ ,  $\theta \in [\underline{\theta}, \infty)$ . Then  $p(h) < \hat{v}$  implies every individual with economic type  $\theta \in [\underline{\theta}, \hat{\theta})$  is rejected by the group. However, if  $\sigma(\underline{\theta}, h) = \sigma(\underline{\theta}, l) = 0$  then  $\tau = (\underline{\theta}, l)$  can profitably deviate to education level  $s_0^c(\underline{\theta}) > 0$ : this individual is at worst rejected in either case and, since  $\underline{\theta}$  is the lowest possible economic type, his or her net income is surely higher with  $s_0^c(\underline{\theta})$  than with no education. So assume  $\sigma(\underline{\theta}, h) = \sigma(\underline{\theta}, l) > 0$  and, first, suppose there exists an interval of economic types  $[\underline{\theta}, \hat{\theta})$  with  $\sigma(\theta, \gamma) = \sigma(\underline{\theta}, l)$  for all  $\gamma$  and all  $\theta \in [\underline{\theta}, \hat{\theta})$ ; then, since  $\sigma$  is an equilibrium strategy,  $\hat{\theta}$  finite and Lemma I imply  $\lim_{\eta \downarrow 0} \sigma(\hat{\theta} + \eta, \gamma) > \sigma(\underline{\theta}, l)$ . Consider a deviation by  $\tau = (\hat{\theta}, \gamma)$  to an education level  $s = \sigma(\underline{\theta}, l) + \epsilon$ , for sufficiently small  $\epsilon > 0$ . Given the maintained assumptions on payoffs (3), D1 yields that both the firms and the group believe the deviant's type is  $(\hat{\theta}, l)$ . To check this, suppose the group believes the deviant's social type is  $\gamma = l$ ; then the individual is still rejected with the deviation and  $\hat{\theta}$  is uniquely the economic type who requires the lowest change in wage to make him or her indifferent between being rejected with the deviation and the candidate equilibrium payoff in which  $(\hat{\theta}, \gamma)$  is rejected by the group with education  $\sigma(\underline{\theta}, l)$ ; on the other hand, if the group believes the deviant's social type is  $\gamma = h$  then it accepts the individual and both  $(\hat{\theta}, l)$  and  $(\hat{\theta}, h)$  have a strict incentive to deviate to  $s$ . Thus  $(\hat{\theta}, l)$  is the type most likely to defect from  $\sigma(\underline{\theta}, l)$  to  $s$ . Hence,  $\sigma$  cannot be pooling in  $\theta \in [\underline{\theta}, \hat{\theta})$  on  $\sigma(\underline{\theta}, l) > 0$ .

Because  $\sigma$  cannot be pooling in  $\theta \in [\underline{\theta}, \hat{\theta})$ ,  $\sigma(\underline{\theta}, \gamma) > 0$  implies that  $\sigma$  is separating in economic type on some interval  $[\underline{\theta}, \theta')$ ; without loss of generality, assume  $\theta' = \hat{\theta}$ . If  $\sigma(\underline{\theta}, \gamma) < s_0^c(\underline{\theta})$  and  $\sigma$  is separating in  $\theta$  on  $[\underline{\theta}, \hat{\theta})$ , then  $\sigma$  is strictly increasing by Lemma I(I) and  $(\underline{\theta}, l)$  can improve his or her payoff by choosing  $s' = \min\{s_0^c(\underline{\theta}), \sigma(\hat{\theta}, \gamma)\}$ . Therefore,  $\sigma(\underline{\theta}, \gamma) \geq s_0^c(\underline{\theta})$ . But then, under the D1 refinement,  $\tau = (\underline{\theta}, h)$  can profitably deviate to education level  $s = 0$ . To see this, first note that  $(\underline{\theta}, h)$  strictly prefers to be accepted with  $s_1^c(\underline{\theta}, h) = 0$  and zero income, to being rejected with  $\sigma(\underline{\theta}, h) > 0$  and positive income. And second, given the deviation from  $\sigma(\underline{\theta}, h)$  to  $s = 0$ , under D1 the group surely believes the deviant is a high social type and so accepts him or her: from (8),  $(\underline{\theta}, l)$  prefers to separate in  $\gamma$  with education  $s_0^c(\underline{\theta})$  to pooling in  $\gamma$  on  $s_1^c(\underline{\theta}, h) = 0$ . Since there are no other possibilities, this proves the lemma.  $\square$

**Lemma IV:** Let  $(\sigma, \alpha)$  be D1 equilibrium strategies. Let  $\sigma_1$  and  $\sigma_0$ , respectively, denote the restriction of  $\sigma$  to those types accepted and rejected by the group in the equilibrium. Then

- (I)  $\sigma_0$  solves the differential equation (9) with  $\alpha = 0$  and initial condition  $\lim_{\theta \downarrow \theta^\circ} \sigma_0(\theta, \gamma) = s_0^c(\theta)$ , where  $\theta^\circ$  is the infimal economic type rejected under  $(\sigma, \alpha)$ ;
- (II)  $\alpha(\sigma_1(\theta, h)) = 1$  and  $\alpha(\sigma_0(\theta, l)) = 0$  implies  $\sigma_1(\theta, h) < \sigma_0(\theta, l)$ .

**Proof.** (I) Since social type is irrelevant conditional on being rejected, rejected individuals face a classical costly signaling problem. Hence, (1) follows directly from Riley [1979].

(II) By Lemma III, the claim is surely true for  $(\underline{\theta}, \gamma)$ . Suppose  $\sigma_1(\theta, h) \geq \sigma_0(\theta, l)$  for some  $\theta > \underline{\theta}$ . Then there exists  $\theta' \in (\underline{\theta}, \theta]$  such that

$$\lim_{\epsilon \rightarrow 0} \sigma_0(\theta' - \epsilon, l) \leq \sigma_1(\theta', h).$$

However, incentive compatibility requires that for all  $\epsilon > 0$ ,

$$u(\omega(\sigma_0(\theta' - \epsilon, l)), 0, \sigma_0(\theta' - \epsilon, l), (\theta' - \epsilon, l)) \geq u(\omega(\sigma'_1), 1, \sigma'_1, (\theta' - \epsilon, l)),$$

where  $\sigma'_1 \equiv \sigma_1(\theta', h)$ . Taking limits as  $\epsilon \rightarrow 0$  yields

$$\lim_{\epsilon \rightarrow 0} u(\omega(\sigma_0(\theta' - \epsilon, l)), 0, \sigma_0(\theta' - \epsilon, l), (\theta', l)) \geq u(\omega(\sigma'_1), 1, \sigma'_1, (\theta', l)),$$

which, by (4), implies

$$\lim_{\epsilon \rightarrow 0} \omega(\sigma_0(\theta' - \epsilon, l)) > \omega(\sigma'_1).$$

By claim (I), for all rejected types  $(\theta, \gamma)$ ,  $\omega(\sigma_0(\theta, \gamma)) = \sigma_0(\theta, \gamma)\theta$ . Hence the preceding inequality is possible only if  $\lim_{\epsilon \rightarrow 0} \sigma_0(\theta' - \epsilon, l) > \sigma'_1 \equiv \sigma_1(\theta', h)$ : contradiction. This proves (II).  $\square$

**Lemma V:** Let  $(\sigma, \alpha)$  be D1 equilibrium strategies. Then there exists some  $\hat{\theta} > \underline{\theta}$  such that  $\alpha(\sigma(\tau)) = 1$  if and only if  $\tau \in [\underline{\theta}, \hat{\theta}) \times \{h\}$ .

**Proof.** Follows directly from Lemmas I, III and IV.  $\square$

Recall that  $\sigma_1$  and  $\sigma_0$ , respectively, denote the equilibrium strategies of those accepted and rejected by the group in the equilibrium.

**Lemma VI:** Let  $(\sigma, \alpha)$  be D1 equilibrium strategies such that, for some  $\hat{\theta} \in (\underline{\theta}, \infty)$  and all  $\epsilon > 0$ ,  $\sigma(\hat{\theta} - \epsilon, l) \neq \sigma(\hat{\theta} - \epsilon, h)$  and  $\sigma(\hat{\theta} + \epsilon, l) = \sigma(\hat{\theta} + \epsilon, h)$ . Then  $\sigma$  cannot be separating in  $\theta$  at  $\hat{\theta}$ .

**Proof.** By Lemma IV(I),  $\sigma_0$  is separating in  $\theta$  and, by Lemma III,  $\sigma$  is separating in  $\gamma$  on some interval  $[\underline{\theta}, \hat{\theta})$ . Suppose, contrary to the claim, that  $\sigma$  is separating in  $\theta$  on  $[\underline{\theta}, \infty)$ . Since  $\sigma$  is presumed an equilibrium strategy and payoffs are continuous in education for each  $\alpha \in \{0, 1\}$ , if  $\hat{\theta}$  is finite then the type  $(\hat{\theta}, h)$  must be indifferent between choosing education  $\sigma_1(\hat{\theta}, h)$  and being accepted by the group, and choosing education  $\sigma_0(\hat{\theta}, h)$  and being rejected by the group. Specifically,

$$(11) \quad u(\omega(\hat{\sigma}_0), 0, \hat{\sigma}_0, (\hat{\theta}, h)) = u(\omega(\hat{\sigma}_1), 1, \hat{\sigma}_1, (\hat{\theta}, h)),$$

where  $\hat{\sigma}_\alpha \equiv \sigma_\alpha(\hat{\theta}, h)$ ,  $\alpha \in \{0, 1\}$ . By definition of  $s_1^c(\hat{\theta}, h)$  and Lemma II,  $\sigma$  separating in  $\theta$  implies

$$(12) \quad u(\omega(\hat{\sigma}_1), 1, \hat{\sigma}_1, (\hat{\theta}, h)) < u(w(s_1^c(\hat{\theta}, h), \hat{\theta}), 1, s_1^c(\hat{\theta}, h), (\hat{\theta}, h)).$$

And again because  $\sigma$  is separating in  $\theta$ ,  $\omega(\sigma(s, \hat{\theta})) = \hat{\theta}s$ ; therefore, by (3),  $u$  is strictly concave in  $s$  and  $u(\omega(s), 0, s, (\hat{\theta}, h)) < u(\omega(s), 1, s, (\hat{\theta}, h))$  for all  $s$ . But then (11) and (12) imply that either  $\hat{\sigma}_1 < s_1^c(\hat{\theta})$  which contradicts Lemma II, or  $\hat{\sigma}_1 > \hat{\sigma}_0$  which contradicts Lemma IV(II): see Figure IV. Therefore  $\sigma$  cannot be separating in  $\theta$  at  $\hat{\theta}$ , proving the claim.  $\square$

Figure IV here

**Proof of Proposition II.** By Lemma V, if  $\sigma$  is a D1 equilibrium strategy, then there exists  $\hat{\theta} > \underline{\theta}$  such that  $\sigma$  is separating in  $\gamma$  for  $\tau = (\theta, \gamma)$ ,  $\theta \in [\underline{\theta}, \hat{\theta})$ . In particular, for  $\theta \in [\underline{\theta}, \hat{\theta})$  the group accepts only high social types  $(\theta, h)$ . By Lemma VI, if types  $\tau = (\theta, \gamma)$ ,  $\theta \in (\hat{\theta}, \theta'')$ , pool in  $\gamma$  (that is,  $\sigma(\theta, h) = \sigma(\theta, l)$ ), then there exists an economic type  $\theta' < \hat{\theta}$  such that  $\sigma(\theta, h) = \hat{s}$  for all  $\theta \in (\theta', \hat{\theta})$ . Moreover, since  $p(h) < \hat{v}$ , all  $\tau$  with  $\theta \in [\hat{\theta}, \theta'')$  are rejected by the group. By Lemma IV(I),  $\sigma_0$  is strictly increasing in  $\theta$  on  $[\underline{\theta}, \theta'')$  and  $\sigma(\underline{\theta}, l) = s_0^c(\underline{\theta}) > 0$ .

With the preceding observations understood, suppose that  $\hat{s} = 0$ . By construction,  $\hat{s} = s_1^c(\underline{\theta}, h) = 0$  and, from the preceding,  $\sigma(\underline{\theta}, l) = s_0^c(\underline{\theta}) > 0$ ; hence (8) implies

$$u(w(s_1^c(\underline{\theta}, h), \underline{\theta}), 1, s_1^c(\underline{\theta}, h), (\underline{\theta}, l)) = 1 + l < u(w(s_0^c(\underline{\theta}), \underline{\theta}), 0, s_0^c(\underline{\theta}), (\underline{\theta}, l)).$$

Similarly,

$$\begin{aligned} 1 + h &= u(w(s_1^c(\underline{\theta}, h), \underline{\theta}), 1, s_1^c(\underline{\theta}, h), (\underline{\theta}, h)) \\ &> u(w(s_0^c(\underline{\theta}), \underline{\theta}), 0, s_0^c(\underline{\theta}), (\underline{\theta}, h)) \\ &= u(w(s_0^c(\underline{\theta}), \underline{\theta}), 0, s_0^c(\underline{\theta}), (\underline{\theta}, l)), \end{aligned}$$

where the inequality follows from (4). Therefore, by continuity, there exists  $\theta_H \in (\underline{\theta}, \infty)$  such that

$$(13) \quad \begin{aligned} 1 + h &= u(\omega(\hat{s}), 1, \hat{s}, (\theta_H, h)) \\ &= u(\omega(\sigma(\theta_H, h)), 0, \sigma(\theta_H, h), (\theta_H, h)) \end{aligned}$$

and, for all  $\theta, \theta'''$  with  $\theta < \theta_H < \theta'''$ ,

$$u(\omega(\sigma(\theta, h)), 0, \sigma(\theta, h), (\theta, h)) < 1 + h < u(\omega(\sigma(\theta''', h)), 0, \sigma(\theta''', h), (\theta''', h)).$$

Let  $\sigma$  be as described in the proposition with  $\underline{\theta} = \theta' = \theta_L$ ,  $\theta'' = \infty$  and  $\hat{\theta} = \theta_H$  defined by (13). Then the group's best response is to accept only those individuals with education  $s \leq \sigma(\theta_H, h) = \hat{s}$ . These strategies are supported as a D1 equilibrium by the group believing with probability one that any deviation to an education level  $s \in (0, s_0^c(\underline{\theta}))$  is chosen by the type  $(\theta, l)$  and best responding by rejecting the individual. This proves there exists a D1 equilibrium as required.

Now suppose  $\hat{s} > 0$ . Then  $\hat{s} < \tilde{s}(\underline{\theta})$ , for otherwise the lowest rejected type,  $(\underline{\theta}, l)$ , weakly prefers choosing  $\tilde{s}(\underline{\theta})$  and being accepted to choosing  $s_0^c(\underline{\theta})$  and being rejected. By Lemma III,  $\sigma(\underline{\theta}, h) = 0$ . There are two possibilities. First, there exists some  $\theta'' > \underline{\theta}$  strictly interior to the set of accepted economic types such that types in an interval  $(\theta, \theta'')$  pool on an education level  $s$ , and types in an interval  $(\theta'', \theta''')$  choose education levels strictly greater than  $s$ ; that is, there exists  $\theta'' > \underline{\theta}$  and some  $\eta > 0$  such that, for all  $\epsilon \in (0, \eta)$ ,

$$0 \leq \sigma(\theta'' - \epsilon, h) = s < \lim_{\theta \downarrow \theta''} \sigma(\theta, h) \leq \hat{s}.$$

But in this instance, D1 implies that any deviation to an education  $s' = s + \delta$ ,  $\delta > 0$  sufficiently small, induces both the group and the firm to believe the deviant's type is  $\tau = (\theta'', h)$ : since  $\tau$  is interior to the accepted set of types,  $\tau$  is uniquely the type who requires the lowest change in wage to make him or her indifferent between being accepted with education  $s'$  and the candidate equilibrium payoff. Therefore such a deviation is strictly payoff-improving to  $\theta''$  and the equilibrium breaks down.

By Lemma VI, the remaining possibility is that  $\sigma(\theta, h)$  is strictly increasing on an interval of accepted economic types  $\tau$  such that  $\theta \in [\underline{\theta}, \theta']$ ,  $\theta' < \hat{\theta}$ ; and  $\sigma(\theta, h) = \hat{s}$  for  $\theta \in [\theta', \hat{\theta})$ . Then  $\sigma_1$  is separating in  $\theta$  on  $[\underline{\theta}, \theta']$  in which case  $\sigma_1$  must satisfy (9) with  $\alpha\gamma = h$  and initial condition  $\sigma_1(\underline{\theta}, h) = s_1^c(\underline{\theta}, h) = 0$ ; denote this strategy by  $\sigma_1^*$ . By Lemma II,  $\sigma_1^*(\theta, h) > s_1^c(\theta, h)$ ; hence,  $\lim_{\theta \rightarrow \infty} \sigma_1^*(\theta, h) = 1 > s_0^c(\underline{\theta}) > \hat{s}$ . Therefore,  $\hat{\theta} < \infty$ , in which case there exists a D1 equilibrium with  $\hat{s} > 0$  if and only if there exists an education level  $\hat{s} < \tilde{s}(\underline{\theta})$  and finite economic types

$\theta_H > \theta_L > \underline{\theta}$  such that:

$$(14) \quad u(\omega(\sigma_1^*(\theta_L, h)), 1, \sigma_1^*(\theta_L, h), (\theta_L, h)) = u(\omega(\hat{s}), 1, \hat{s}, (\theta_L, h)),$$

$$(15) \quad u(\omega(\hat{s}), 1, \hat{s}, (\theta_H, h)) = u(\omega(\sigma_0(\theta_H, h)), 0, \sigma_0(\theta_H, h), (\theta_H, h))$$

and

$$(16) \quad \omega(\hat{s}) = \frac{\hat{s}}{F(\theta_H) - F(\theta_L)} \int_{\theta_L}^{\theta_H} \theta dF(\theta).$$

Given such  $\hat{s}$ ,  $\theta_H$  and  $\theta_L$ , let  $\sigma$  be as described in the proposition . Then the group's best response is to accept only those individuals with education  $s \leq \sigma(\theta_H, h) = \hat{s}$ . These strategies are supported as a D1 equilibrium by the group (and firms) believing with probability one that any deviation to an education level  $s \in (\hat{s}, s_0^c(\underline{\theta}))$  is chosen by the type  $(\theta_H, l)$  and best responding by rejecting the individual; and believing with probability one that a deviation to an education level  $s \in (\sigma_1^*(\theta_L, h), \hat{s})$  is chosen by  $(\theta_L, h)$  (since  $\hat{s} < \tilde{s}(\underline{\theta})$ , no low social type can profit by making such a deviation and, by definition of  $\sigma_1^*$  and (14), no accepted type can profit either). This completes the proof.  $\square$

Department of Managerial Economics and Decision Sciences, Northwestern University  
Harvard University Society of Fellows and NBER

## References

- [1] Ainworth-Darnell, James and Douglas Downey, "Assessing the Oppositional Culture Explanation for Racial/Ethnic Differences in School Performance," *American Sociological Review*, LXIII (1998), 536-553.
- [2] Akerlof, George, "The Economics of the Rat Race and Other Woeful Tales," *Quarterly Journal Economics*, XC (1976), 599-617.
- [3] Akerlof, George, "A Theory of Social Custom of Which Unemployment May Be One Consequence," *Quarterly Journal Economics*, XCIV (1980), 599-617.
- [4] Akerlof, George and Rachel Kranton, "Economics and Identity." *Quarterly Journal of Economics*, CXV (2000), 715-753.

- [5] Anderson, Elijah, *Streetwise: Race, Class, and Change in an Urban Community*, (Chicago, IL: University of Chicago Press, 1990).
- [6] Anderson, Elijah, *Code of the Streets: Decency, Violence, and the Moral Life of the Inner City*, (Norton, W. W. & Company, Inc., 1999).
- [7] Austen-Smith, David, "Peer Pressure and Job Market Signalling," *Center for Mathematical Studies in Economics and Management Science*, Discussion Paper No. 1352, 2002, Northwestern University.
- [8] Austen-Smith, David, and Roland G. Fryer, Jr, "The Economics of 'Acting White.'" *NBER Working Paper No. 9904*, 2003.
- [9] Banks, Jeffrey, and Joel Sobel, "Equilibrium Selection in Signaling Games," *Econometrica*, LV (1987), 647-661.
- [10] Bernheim, Douglas, "A Theory of Conformity," *Journal of Political Economy*, CII (1994), 841-877.
- [11] Bowles, Samuel and Herbert Gintis, "Persistent Parochialism: Trust and Exclusion in Ethnic Networks," *Journal of Economic Behavior and Organization* LV, (2004), 1-23
- [12] Brooks-Gunn, Jeanne and Greg J. Duncan, eds., *The Consequences of Growing Up Poor*, (New York: Russell Sage, 1997).
- [13] Brown, Charles, "Relatively Equal Opportunity in the Armed Forces: Impacts on Children of Military Families," work in progress, University of Michigan, 2000.
- [14] Calvo-Armengol, Toni. and Matthew Jackson, "Social Networks in Determining Employment: Patterns, Dynamics, and Inequality," *American Economic Review*, XCIV (2004), 426-444.
- [15] Case, Anne, and Lawrence Katz, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." *NBER Working Paper No. 3705*, 1991.
- [16] Chapple, Simon, Richard Jefferies, and Rita Walker, *Maori Participation and Performance in Education: A Literature Review and Research Programme*. Report for the Ministry of Education, (1997).

- [17] Chin, Tiffani and Meredith Phillips, "Oppositional to What? Achievement Ideologies, Resistance, and Ethnic Authenticity among Urban Youth." Mimeo. University of California, Los Angeles, 2004.
- [18] Cho, In-Koo and David Kreps, "Signaling Games and Stable Equilibria," *Quarterly Journal Economics*, CII (1987), 179-221.
- [19] Cho, In-Koo and Joel Sobel, "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, L (1990), 381-413.
- [20] Cole, Harold, George Mailath and Andrew Postlewaite, "Social Norms, Savings Behavior and Growth," *Journal of Political Economy*, C (1992), 1092-1225.
- [21] Coleman, James, and others, *Equality of Educational Opportunity*, Washington DC: U.S. Government Printing Office, (1966).
- [22] Cook, Phillip and Jens Ludwig, "The Burden of 'Acting White': Do Black Adolescents Disparage Academic Achievement?" In Christopher Jenks and Meredith Phillips, eds., *The Black-White Test Score Gap*. (Brookings Press, 1998).
- [23] Corwin, Miles, *And Still We Rise: The Trials and Triumphs of Twelve Gifted Inner-City Students*, (New York, NY: Harper Collins, 2001).
- [24] Cutler, David. and Glaeser, Edward, "Are Ghettos Good or Bad?," *Quarterly Journal of Economics*, CXII (1997) , 827-872.
- [25] Devos, George and Hiroshi Wagasutma, *Japan's Invisible Race: Caste in Culture and Personality*, (Berkeley: University of California Press, 1966).
- [26] Drake, St. Clair and Horace Cayton, *Black Metropolis: A Study of Negro Life in a Northern City*, (Chicago: University of Chicago Press, 1945).
- [27] Ferguson, Ronald, "A Diagnostic Analysis of Black-White GPA Disparities in Shaker Heights, Ohio." *Brooking Papers on Education Policy*, (2001), 347-414.
- [28] Ferguson, Ronald, "What Doesn't Meet the Eye: Understanding Racial Disparities in Fifteen Suburban School Districts," mimeo, Harvard University, 2002.

- [29] Fordham, Signithia, *Blacked Out: Dilemmas of Race, Identity, and Success at Capital High*, (Chicago: University of Chicago Press, 1996).
- [30] Fordham, Signithia and John Ogbu, Black students' school success: coping with the "Burden of Acting White," *The Urban Review*, XVIII (1986), 176-206.
- [31] Freedman, Jonathan, Merrill Carlsmith, and David Sears, *Social Psychology*, (Englewood Cliffs, NJ: Prentice Hall, 1970).
- [32] Fryer, Roland, "The Economics of 'Acting White'," mimeo, University of Chicago, 2001.
- [33] Fryer, Roland, "An Economic Approach to Cultural Capital." manuscript, Harvard University, 2004.
- [34] Fryer, Roland and Paul Torelli, "Understanding the Prevalence and Impact of 'Acting White'," Working paper, Harvard University, 2004.
- [35] Gans, Herbert, *The Urban Villagers: Group and Class in the Life of Italian-Americans*, (New York, NY: The Free Press, 1962).
- [36] Jargowsky, Paul, and Bane, Mary Jo., "Ghetto Poverty in the United States, 1970-1980." in *The Urban Underclass*, Christopher Jencks, and Paul Peterson eds., The Brookings Institution: Washington D.C., 1991.
- [37] Jensen, Arthur, *The G Factor: The Science of Mental Ability*, (Westport, CT: Praeger Publishers, 1998).
- [38] Kain, John, "Housing Segregation, Negro Employment, and Metropolitan Decentralization," *Quarterly Journal of Economics*, LXXXII, (1968), 175-197.
- [39] Kling, Jeffrey R., and Jeffrey B. Liebman. "Experimental Analysis of Neighborhood Effects on Youth." *Princeton IRS Working Paper No. 483*, (2004).
- [40] Lindbeck, Assar, Sten Nyberg, and Jorgen Weibull, "Social Norms and Economic Incentives in the Welfare State," *Quarterly Journal of Economics*, CXIV (1999), 1-35.
- [41] Massey, Douglas and Nancy Denton, *American Apartheid: Segregation and the Making of the Underclass*, (Cambridge, MA: Harvard University Press, 1993).

- [42] McWhorter, John, *Losing the Race: Self-Sabotage in Black America*. (New York, NY: Free Press, 2000).
- [43] Ogbu, John and Astrid Davis, *Black American Students in an Affluent Suburb: A Study of Academic Disengagement*, (Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 2003).
- [44] Olson, Steve, *Mapping Human History: Genes, Race, and Our Common Origins*. (Houghton Mifflin Company, 2002).
- [45] Pattillo-McCoy, Mary, *Black Picket Fences: Privilege and Peril Among the Black Middle Class*, (Chicago: University of Chicago Press, 1999).
- [46] Riley, John, "Informational Equilibrium," *Econometrica*, XLVII (1979), 331-360.
- [47] Rubinstein, Yona. "The Slowdown in the Black-White Wage Convergence: An Homogenous Phenomenon." mimeo. Tel Aviv University, (2001).
- [48] Rushton, Phillip, *Race, Evolution, and Behavior: A Life History Perspective*, (New Brunswick, NJ: Transaction Publishers, 2000).
- [49] Smrekar, Claire, and others, "March Toward Success: School Success and Minority Student Achievement in Department of Defense Schools." (2001).
- [50] Spence, Michael, *Market Signaling*, (Cambridge: Harvard University Press, 1974).
- [51] Suskind, Ronald, *A Hope in the Unseen: An American Odyssey From the Inner City to the Ivy League*, (New York: Broadway, 1998).
- [52] Sykes, Bryan, *The Seven Daughters of Eve: The Science That Reveals our Genetic Ancestry*, (New York, NY: W.W. Norton Company, 2002).
- [53] Tyson, Karolyn, William Darity, Jr. and Domini Castellino, "Breeding Animosity: The "Burden of Acting White" and Other Problems of Status Group Hierarchies in Schools." Working Paper, University of North Carolina at Chapel Hill, 2004.
- [54] Van Deburg, Willia, *New Day in Babylon: The Black Power Movement and American Culture, 1965-1975*. Chicago: University of Chicago Press.
- [55] Willis, Paul, *Learning to Labor: How Working Class Kids Get Working Class Jobs*, (New York, NY: ColumbiaUniversity Press, 1977).

- [56] Wilson, William-Julius, *The Declining Significance of Race: Blacks and Changing American Institutions*, (The University of Chicago Press, 1978).

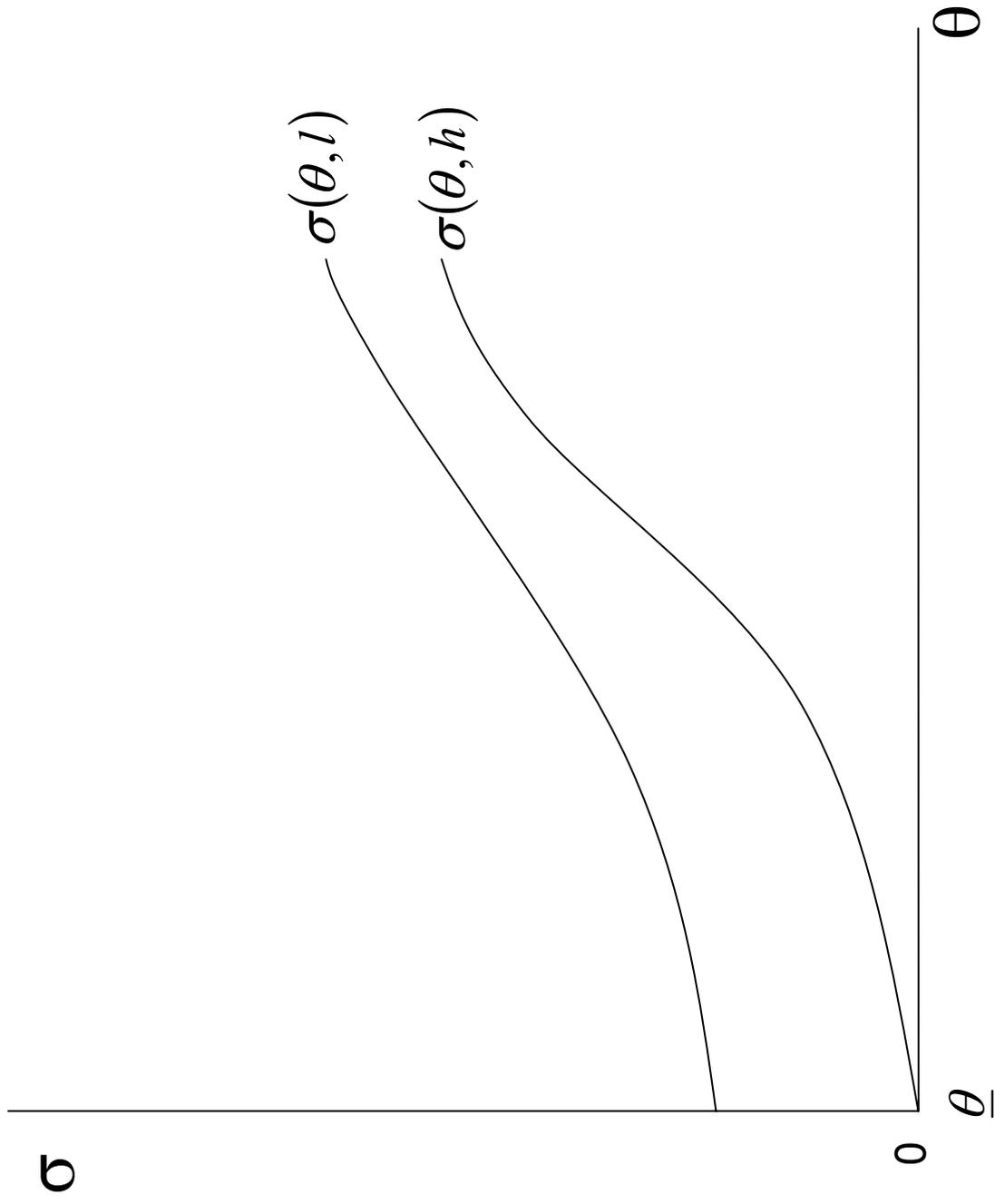


Figure I: Separating in  $\gamma$ ,  $\theta$  common knowledge

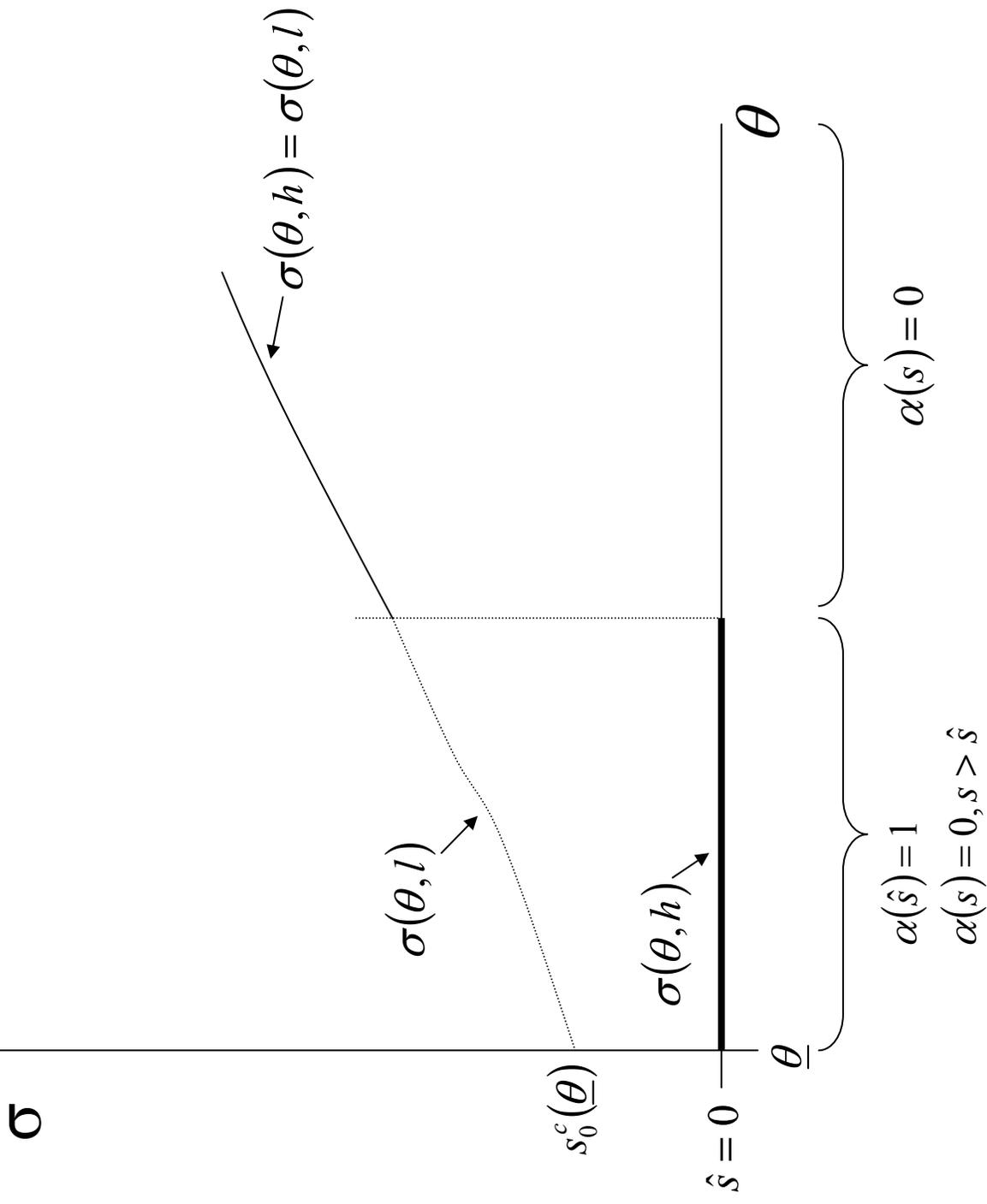


Figure IIA: all accepted types pool on minimal education

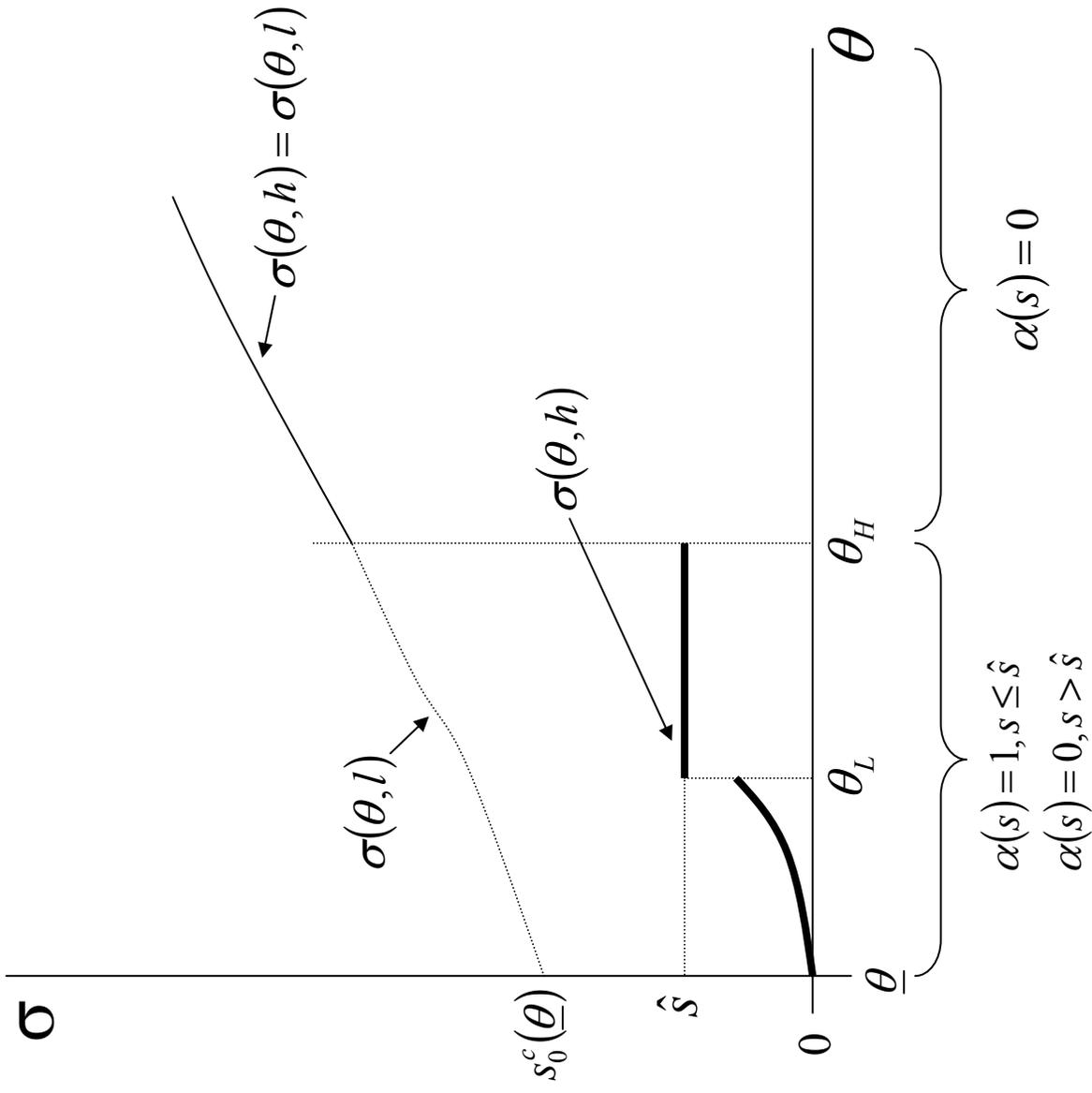


Figure IIB: highest accepted types pool on positive education

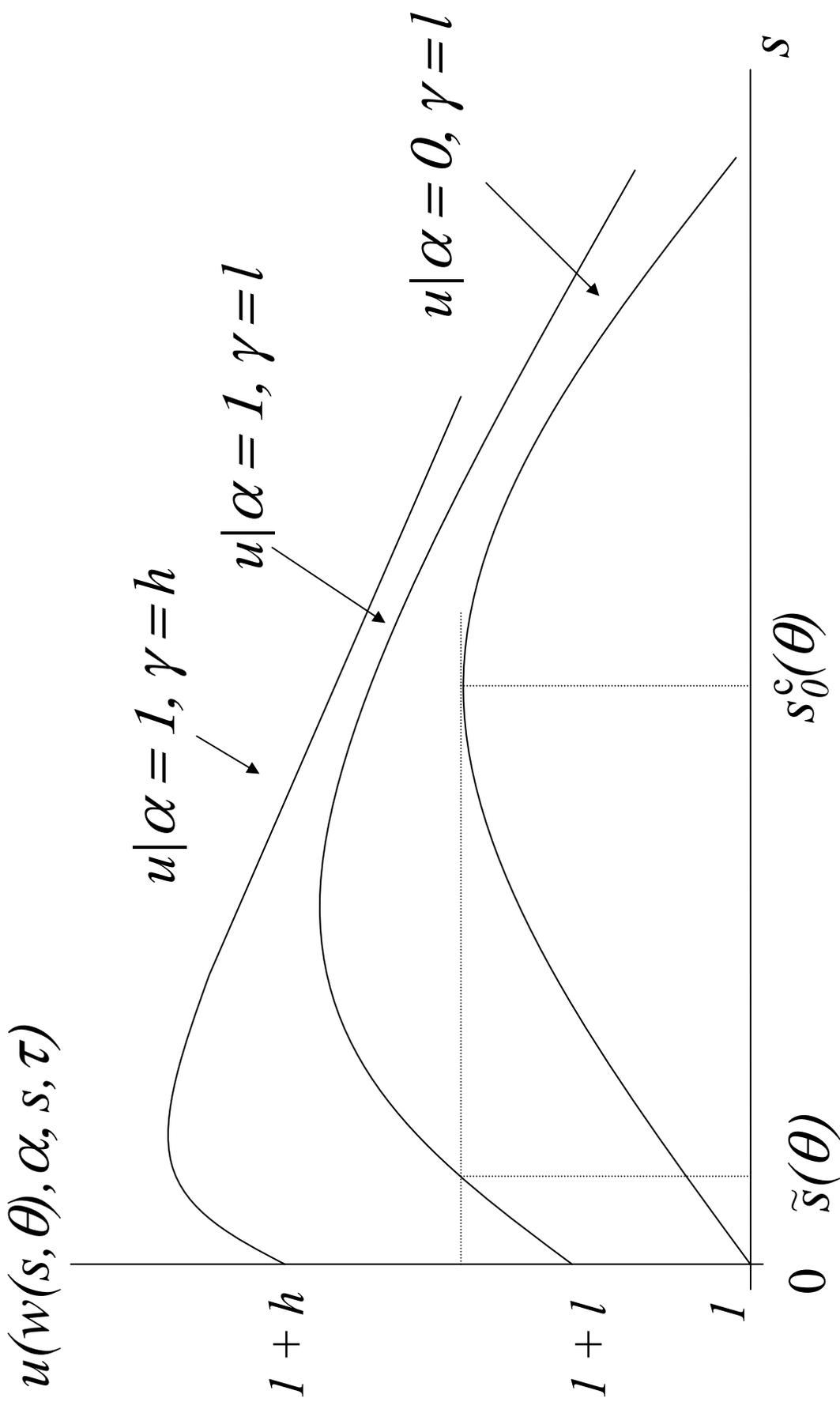


Figure III: complete information payoffs for  $\tau = (\theta, \gamma)$

$$u(w(s, \hat{\theta}), \alpha, s, (\hat{\theta}, h))$$

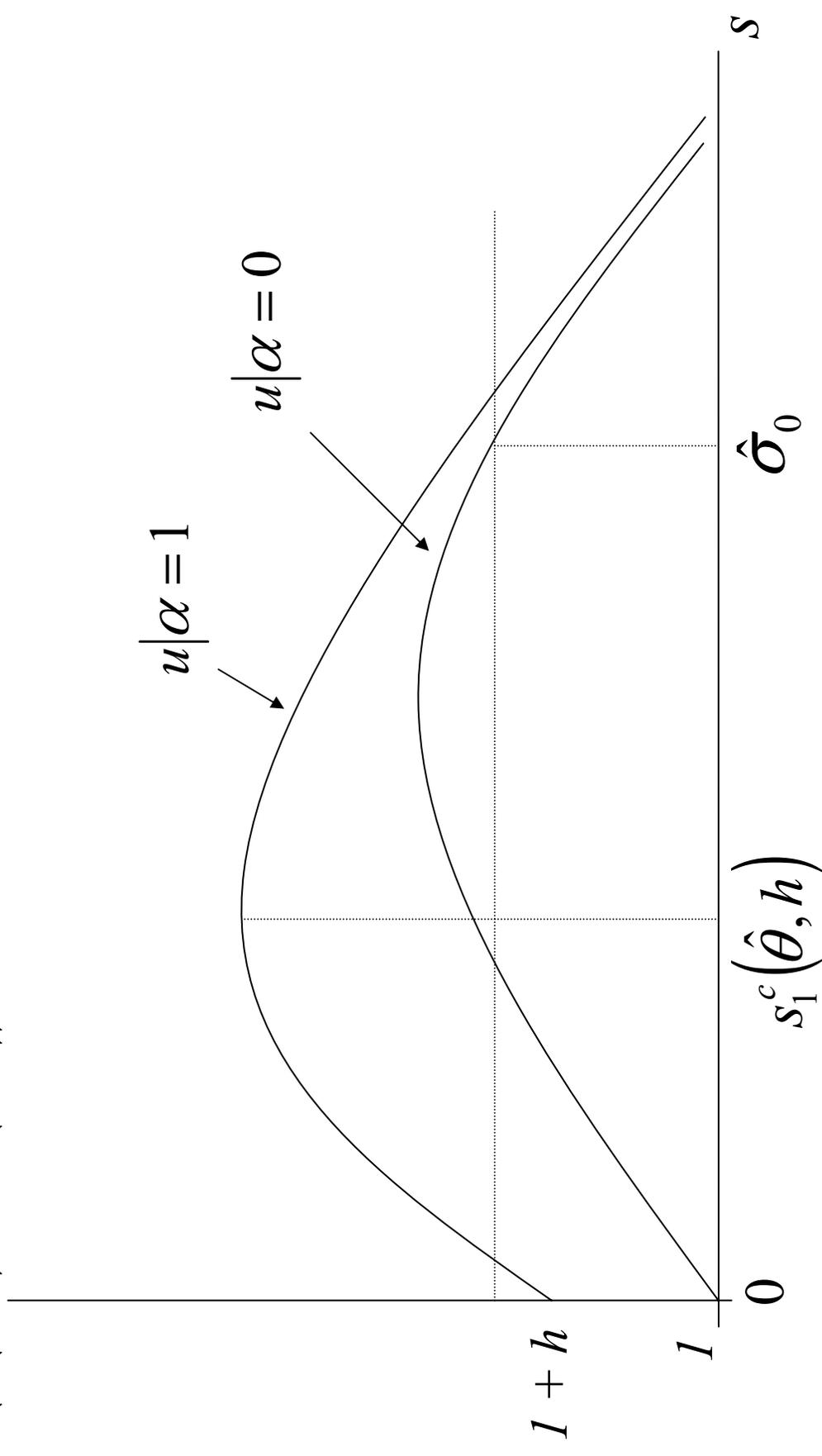


Figure IV: complete information payoffs for marginal type