

JOB DISCRIMINATION, MARKET FORCES
AND THE INVISIBILITY HYPOTHESIS

CFDP No 708 R

Paul Milgrom* and Sharon Oster**
Yale University, Box 1A
New Haven, CT. 06520

June, 1984
revised June, 1985

* School of Organization and Management, Economics Department, and Cowles Foundation, Yale University.

* School of Organization and Management and Economics Department, Yale University.

Acknowledgements. We would like to thank Rick Levin, Asher Wolinsky, Drew Fudenberg and the participants of the Yale Microeconomics Seminar and the Cowles seminar for their helpful comments, and Kurt Anstreicher for his help with the technical appendix. Milgrom's work was supported by NSF grant IST-8411595 to Yale and by the Institute for Advanced Studies of the Hebrew University of Jerusalem.

ABSTRACT

The Invisibility Hypothesis holds that the job skills of disadvantaged workers are not easily discovered by potential new employers, but that promotion enhances visibility and alleviates this problem. Then, at a competitive labor market equilibrium, firms profit by hiding talented disadvantaged workers in low level jobs. Consequently, those workers are paid less on average and promoted less often than others with the same education and ability. As a result of the inefficient and discriminatory wage and promotion policies, disadvantaged workers experience lower returns to investments in human capital than other workers.

Key Words: Discrimination, Human Capital, Job Discrimination, Labor Markets, Visibility.

Journal of Economic Literature Classifications:

Primary - 821

Secondary - 026, 826

JOB DISCRIMINATION, MARKET FORCES
AND THE INVISIBILITY HYPOTHESIS

by Paul Milgrom and Sharon Oster

Among non-economists, job discrimination is rarely regarded as a phenomenon that can be understood in purely economic terms, without reference to its richer social context. More often, it is seen as just one thread in a fabric of discrimination encompassing housing, education, jobs, club memberships, and more.

In contrast, textbook economic theory holds that a worker's pay and chances for promotion depend *only* on his or her productive abilities. The presence of prejudiced employers who discriminate against, say, non-white workers or women is irrelevant according to this view: The remaining employers will be led by the pursuit of profits to hire the underpaid and promote the underemployed. Competition among unprejudiced employers will result in rising wages for victims of discrimination. Therefore, the existence of job discrimination is at worst a minor and temporary phenomenon, or so the usual argument goes.

Given that there is discrimination, which we take to mean that employers pay women and non-whites less (on average) than equally able white males and assign them to less responsible positions, the puzzle is: How can market forces allow such a state of affairs to continue? One possibility, explored by Becker (1957), is that individuals have a taste for discrimination — a willingness to earn lower profits or receive lower wages in order to maintain segregation in the workplace. However, as Arrow (1972) pointed out, in order for a taste for

discrimination to result in lower wages, it must be quite widespread. In an otherwise homogeneous labor market, the targets of discrimination will be employed by the least prejudiced employers — those with the least willingness to sacrifice profits to participate in discrimination — and will work together with the least prejudiced workers. If there are sufficiently many unprejudiced workers and employers, the tastes of the others can lead to segregation in employment, but not to significantly lower wages for the victims of discrimination. Thus, the application of Becker's theory to explain, say, discrimination against blacks in contemporary American society, rests on the hypothesis that the prejudice against blacks is strong both in its extent (proportion of the population who are prejudiced) and in its degree (the compensation they demand to accept integration).

Another view, advanced by Phelps (1972), tries to explain discrimination without resorting to an assumption of prejudice in the workplace. Phelps's theory is based on the hypothesis that traditional indicators of ability are less informative when applied to certain "disadvantaged" workers. Rational employers therefore place less reliance on them and so pay the ablest disadvantaged workers less on average than they pay equally able advantaged workers. However, according to that theory, employers also pay the least able disadvantaged workers *more* than their advantaged counterparts. The "statistical theory of discrimination" does not explain discrimination as we have defined it; it predicts that the average wage paid to members of any group that is distinguished by some *observable* characteristics (such as

sex or race) to depend only on the average productive prowess among the group's members.

A recent paper by Lundberg and Startz (1983) extends the statistical theory of discrimination to include a human capital investment choice by the workers. Like Phelps, Lundberg and Startz assume that each worker is paid his or her expected marginal product, given the available indicators. Then, since disadvantaged workers' job skills are less accurately evaluated than those of other workers, they have less incentive to make productivity-enhancing investments in human capital. Responding to these incentives, they acquire less human capital, and their average productivity and wages are correspondingly lower.

The models that we have so far discussed make no allowance for the fact that firms have a variety of different jobs to be filled, each requiring its own particular set of skills. If, as Phelps and Lundberg and Startz have supposed, employers were less able to evaluate the job skills of women and non-whites, these disadvantaged workers would be less well matched to jobs than their advantaged counterparts. Then, even if the distributions of skills among advantaged and disadvantaged workers were the same, productivity would be lower for the disadvantaged workers. That, alone, would explain the lower wages paid to disadvantaged workers.

This matching effect is distinguished from the other effects described above by the fact that it results from an inefficient use of labor resources by firms. It appears, therefore, that firms will have the usual incentive to improve efficiency by improving their evaluation of workers' skills. To the extent

that firms succeed in their efforts to assign individual workers more efficiently, they will as a by-product also reduce the other effects associated with statistical discrimination.

However, the argument that firms always have an incentive to use workers efficiently itself relies on a suspect assumption — the assumption that wages are unaffected by job assignments. Waldman (1984) has shown that if (i) workers' skills are privately known to firms and (ii) other potential employers use job assignments as a signal of ability, the wage paid to a worker will depend on his or her job. In that case, as Waldman shows, the textbook efficiency argument breaks down.

Our theory emphasizes the role in the job market of a worker's *visibility*, by which we mean roughly the amount of information about the worker's job skills that potential employers possess. In our model, we employ only two visibility classes; *Visibles* are workers whose abilities are known to all potential employers and *Invisibles* are those whose abilities can be concealed by an employer from other potential employers.

Our formal analysis of the problem of job discrimination rests on a two-part assumption that we call the *Invisibility Hypothesis*. The first part is the assumption that disadvantaged workers are relatively more likely to be Invisible when they first enter the labor market. The second part is that a worker's visibility is enhanced by assignment to a higher level job.

The Invisibility Hypothesis reflects the observation that talent is not inevitably and universally recognized, and that those with advantaged backgrounds are more likely to be recognized for their abilities. There are many causes contributing to the relative lack of recognition for disadvantaged workers. Prejudice — in the form of misperceptions rather than antipathy — can cause an employer to overlook a potentially good employee. So, too, can the failure of an employee to "toot his own horn," whether the reluctance to do so comes from shyness or pride or cultural taboos. The existence of clubs which limit the membership of women, non-whites, or religious or ethnic minorities; job segregation which is not *per se* inefficient but which keeps some people out of view; exclusive neighborhoods; out-of-town conventions that are hard for some working mothers to attend — all of these things contribute to a separation that makes some workers less visible to potential new employers.

Lack of visibility leads to wage distortions by impairing the competition among firms for workers. Employers with superior information about their employees' skills may find that they can earn excess profits on abler workers *so long as those workers can be hidden from other potential employers*. Given our hypothesis that promotion enhances visibility, there is a motive for employers to discriminate against talented Invisibles in promotions. The consequent inefficiency depresses average wages most for the disadvantaged groups, where the invisibility problem is greatest. Since the mechanism of discrimination most depresses the wages of the ablest Invisibles, disadvantaged workers are

discouraged from acquiring human capital, which further depresses the group's productivity and average wages.

We investigate these and other implications of the Invisibility Hypothesis using a simple partial equilibrium model in which there are only two types of jobs, low level and high level. We assume that more talented workers are more productive in any kind of job, and have a comparative advantage in higher level jobs. Also, initially, we assume that employment contracts specify a wage and provide a guarantee against wage cuts or layoffs.¹ In our model, employers buy labor at market-clearing prices to maximize profits, and workers invest in education and other forms of human capital to maximize expected wages net of investment costs. As noted previously, we assume that there are two kinds of workers: Visibles, whose abilities are known to all potential employers, and Invisibles, whose abilities are known only to their own current employers. The model leads to a rich set of predictions, including the following ones:

1. Average wages will be lower for Invisibles than for Visibles. (Hence, groups such as women and certain minorities which have a relatively high proportion of Invisibles among their members will have lower average wages.)

2. Wage variability within job classes will be less for Invisibles. (This prediction of our model is not also implied by Becker's "taste for discrimination" model.)

3. Promotions will be less probable for Invisibles. The ablest Invisibles will be inefficiently assigned to low level jobs.

4. No Invisibles will be among the most highly paid workers.

5. The returns to unobservable investments in human capital will be lower for Invisibles.

6. The marginal returns to observable investments in "general education" (education that enhances productivity proportionately in each kind of job) are lower for Invisibles provided that either of the following two conditions hold: (i) workers of below average ability never have a comparative advantage at high level jobs or (ii) investments in general education are predominantly most productive for the ablest workers.

7. As a corollary to 5 and 6, an Invisible will acquire less unobservable human capital and (when the conditions of 6 hold) less general education than an equally able Visible.

8. Education choices of Invisibles and potential Invisibles are distorted. The Invisibles whose abilities would be most enhanced by education may choose to forgo education, while others whose skills would benefit less may acquire more education.

9. If workers know their own abilities, then an education subsidy that led Invisibles to acquire as much education as Visibles would still lead to an inefficient distribution of education among members of the Invisible group. That, together with the job discrimination, would lead to persistently lower wages and lower probabilities of promotion for Invisibles.

10. An affirmative action program that set both quotas and wage standards would, in the short run, have ambiguous effects on efficiency. It would lead to the promotion of some qualified workers who had been inefficiently assigned to low level jobs, but it would also result in the promotion of some unqualified Invisibles, who would be overpaid relative to their productivity. In the long run, if the quotas and wage standards were set correctly, such an affirmative action program would improve the efficiency of both job assignments and human capital acquisition decisions and would reduce differences in wages and assignments across groups. (There is a corresponding market equilibrium in "employment contracts" that Pareto dominates the equilibrium we have identified, but no *single* firm can benefit by being the first to offer such an employment contract, so these contracts do not overturn the non-contracting equilibrium.)

Because our model incorporates all of the usual labor market forces contained in standard competitive theory and still leads to discrimination, it establishes that the effects of segregation and prejudice, as embodied in the Invisibility Hypothesis, can reduce — and even twist — the propensity of market forces to eliminate discrimination. Moreover, although our theory identifies employers with firms, the analysis need not stop there: invisibility can be a problem even within a corporation. In an

analysis closely akin to ours, Kanter (1977) has identified a phenomenon that she calls the "dilemma of indispensability," in which bosses try to conceal able secretaries from others in the same firm who might bid them away. Secretaries are trapped in lower level jobs by their own talents — their own indispensability to their bosses, and by their invisibility.

THE MODEL

We analyze a simple model, in which all firms are identical and have production functions $f(x,y)$ with two inputs: labor in low level jobs x and labor in high level jobs y . Employment contracts specify a wage $w(e,\eta)$ for those with education e and known ability η and a wage $w(e)$ for those of unknown ability. Wage reductions and layoffs are prohibited. Workers of known ability consist of the initial group of Visible plus any initially Invisible workers who have been promoted.

The two labor inputs, x and y , are measured in efficiency units. A worker of native ability η who has acquired education e provides $\alpha(e,\eta)$ efficiency units of low level labor or $\alpha(e,\eta)\beta(e,\eta)$ units of high level labor, depending on his or her job assignment. Note that β is a measure of the worker's comparative efficiency in high level jobs. It is assumed that abler workers are more productive in both kinds of jobs and have a comparative advantage in high level jobs. Thus, η is a real variable and both α and β are increasing in η .

Our technical assumptions are these: The worker productivity functions α and β are positive, bounded, and continuously differentiable with $\partial\alpha/\partial\eta > 0$ and $\partial\beta/\partial\eta > 0$. The production function f is a smooth neoclassical production function: it is twice continuously differentiable with constant returns to scale and a diminishing marginal rate of substitution among its two inputs. Positive amounts of both kinds of inputs are required for production. To keep notation to a minimum, f is measured in revenue units.

Labor is assumed to be supplied inelastically, so the numbers N_V of Visibles and N_I of Invisibles per firm are specified exogenously. Also, at the time wages are determined, the human capital acquisition decisions are assumed to have already been made, so the distribution of education and talents among workers are parameters from the firms' point of view. Let the joint density of education and ability among Visibles be $g(e,\eta)$ and among Invisibles be $h(e,\eta)$. Finally, we assume that the distribution of abilities given education among the initially Invisible workers is the same for any one firm as for the population at large.

The firm's problem is to choose how many workers of each type to hire, and how to assign workers to jobs. Let $n(e,\eta)$ designate the number of Visibles with education e and ability η that the firm hires, and let $m(e)$ be the number of Invisibles hired with education e . Let $p(e,\eta)$ and $q(e,\eta)$ be the proportions of Visibles and Invisibles, respectively, with education e and ability η that are assigned to high level jobs. The functions n , m , p , and q are the employer's choice variables. These choices

determine the quantities of efficiency labor $x = x_V + x_I$ and $y = y_V + y_I$ and the total wage bill $W = W_V + W_L + W_H$ by the equations:

$$(1a) \quad x_V = \iint \alpha(e, \eta) n(e, \eta) [1 - p(e, \eta)] d\eta de$$

$$(1b) \quad x_I = \iint \alpha(e, \eta) m(e) h(\eta|e) [1 - q(e, \eta)] d\eta de$$

$$(1c) \quad y_V = \iint \alpha(e, \eta) \beta(e, \eta) n(e, \eta) p(e, \eta) d\eta de$$

$$(1d) \quad y_I = \iint \alpha(e, \eta) \beta(e, \eta) m(e) h(\eta|e) q(e, \eta) d\eta de$$

$$(2a) \quad W_V = \iint w(e, \eta) n(e, \eta) d\eta de$$

$$(2b) \quad W_L = \iint w(e) m(e) h(\eta|e) [1 - q(e, \eta)] d\eta de$$

$$(2c) \quad W_H = \iint \max\{w(e), w(e, \eta)\} m(e) h(\eta|e) q(e, \eta) d\eta de.$$

The labor inputs for the firm consist of the low level efficiency units x and high level units y provided by Invisibles and Visibles, as determined by the hiring and job assignment policies. Visibles must be paid their market wage, regardless of the job assignment. Invisibles assigned to low level jobs are paid their entry wages. If an Invisible is assigned to a high level job and if his or her market wage exceeds the entry wage, then the higher market wage must be paid to prevent the worker from being bid away. All of these facts are reflected in the set of equations above.

The labor market is in short-run equilibrium when, at specified wages, markets clear. This means first of all that the number of Visibles demanded at each education-and-ability combination and the optimal number of Invisibles demanded at each education level are equal to their respective supplies. Mathematically, these market-clearing conditions take the form:

$$(3a) \quad n(e, n) = N_V g(e, n)$$

$$(3b) \quad m(e) = N_I h(e) .$$

In addition, it must never be profitable for a firm to raid another firm in an attempt to hire away one of its employees. For Visible workers, this is no additional condition at all; it is implied by the first market clearing condition. For initially Invisible workers who have become Visible through promotion, the "no raids" condition is implicit in expression (2c). For initially Invisible workers who are not promoted and remain Invisible, the issues involved in raids are more subtle, because the potential employer has less information about the employee than the firm being raided.

Perhaps the simplest specification would be to require that no raiding firm can ever make an offer to an employee that the employee would accept and that would be profitable for the raider. This would lead to the constraint that the wage paid to unpromoted Invisibles be not less than their average marginal product in the low level job. However, this constraint relies on the extreme assumption that the current employer cannot make a counteroffer. If the employer were allowed to match outside offers, then a raid would be "successful" only when the worker is not worth the wage offered. The raider would suffer an extreme form of what bidding theorists call the "Winner's Curse" (c.f. Milgrom and Weber (1982)): It would acquire workers only when it offered more than they are worth. In such a situation, a raider would never offer more than value of a worker of minimal ability.

The no-raids constraint would then be

$$(3c) \quad w(e, \eta_{\min}) \leq w(e),$$

which is implied by simply market clearing.

Although neither of these two accounts of raids, the one in which the employer gets to make no counteroffer and the one in which the employer gets the final word, is a satisfactory description of reality, it seems that the second account is better, and it is quite simple analytically. Rather than complicate the model with a fuller account of raids, we adopt the no-raids constraint (3c). With this treatment, equilibrium is characterized as follows.

Proposition 1. A necessary and sufficient condition for equilibrium is that markets clear ((3a)-(3b)) and, in addition, the wages and production policies satisfy (4)-(10) below:

$$(4) \quad w(e, \eta) = \alpha(e, \eta) \max(f_1, f_2 \beta(e, \eta))$$

$$(5) \quad w(e) = f_1 \frac{\int \alpha(e, \eta) h(\eta|e) [1 - q(e, \eta)] d\eta}{\int h(\eta|e) [1 - q(e, \eta)] d\eta}$$

$$(6) \quad p(e, \eta) = \begin{cases} 1 & \text{if } \eta > \eta^*(e) \\ 0 & \text{if } \eta < \eta^*(e) \end{cases}$$

$$(7) \quad q(e, \eta) = \begin{cases} 1 & \text{if } \eta^*(e) < \eta < \eta^{**}(e) \\ 0 & \text{otherwise} \end{cases}$$

$$(8) \quad f_1 = \partial f / \partial x \quad \text{and} \quad f_2 = \partial f / \partial y$$

where f_1 and f_2 are the market-determined wage parameters. The breakpoint $\eta^*(e)$ is given by equation (9) and $\eta^{**}(e)$ is given by equation (10).

$$(9) \quad \beta(e, \eta^*(e)) = f_1/f_2$$

$$(10) \quad \alpha(e, \eta^{**}(e)) = w(e)/f_1$$

The essential content of Proposition 1 is seen from a comparison of (6) and (7): Invisibles suffer discrimination in promotions. Visibles assigned to higher level jobs are precisely those who are more productive in those jobs, that is, those for whom $\eta > \eta^*(e)$. An Invisible assigned to a high level job must meet two conditions. First, like Visibles, he or she must be more productive in that job than in a low level job. Second, his or her marginal product in the low level job must be *less* than the wage paid in that job. Of course, there may be no Invisibles who meet both of these conditions. In that case the Proposition asserts that none are assigned to high level jobs.

Why do firms practice this form of discrimination in which the ablest Invisibles are not promoted? Employers earn rents on the ablest Invisibles in low level jobs, since low level Invisibles are paid the *average* marginal product of all workers in their class, while the ablest individuals, those with $\eta > \eta^{**}(e)$, have marginal products in excess of the class average. Since the ability of a promoted worker becomes known to other potential employers, it is impossible for a firm to earn rents on a high level worker. Therefore, to maximize profits, an employer must inefficiently assign the ablest Invisibles to low level jobs. To do otherwise jeopardizes the rents earned from them.

At equilibrium in a competitive labor market, employers earn zero expected rents on each individual worker hired, although *a posteriori* they may earn positive rents on some workers and negative rents on others. Thus, at equilibrium, the total wage bill paid to initially Invisible workers is just the average marginal product of those workers in the jobs they occupy. Consequently, the loss of efficiency from the improper assignment of Invisible workers falls wholly on the workers themselves, in the form of reduced average wages. We summarize these conclusions in the form of a Proposition.

Proposition 2. Suppose that the joint distributions of education and ability were the same for Visibles and Invisibles. Then the average wage earned by Invisibles would be lower than for Visibles. Moreover, the probability of promotion (that is, of assignment to a high level job) would be lower. If any Invisibles were promoted, the highest wage paid to a Visible would be higher than the highest wage paid to any Invisible.

Now, consider the difference in the wages paid to Visibles and Invisibles as a function of their skill levels. It is entirely possible in this model that no Invisibles satisfy the conditions necessary for promotion. That is, it may occur that $\eta^{**}(e)$ is less than $\eta^*(e)$ for all values of e , in which case all Invisibles who are able enough to be promoted are also sources of rents to their employers. Then, no Invisible workers would be promoted and all would be paid the same entry wage. The difference between the wages paid to Visibles of type (e, η) and Invisibles of the same type would be $w(e, \eta) - w(e)$. That differ-

rence is increasing in η because abler Visible workers command higher wages (as is readily verified from Proposition 1), while Invisibles all earn the same wage.

The other possibility is that some Invisibles (albeit not the most able) are promoted and earn a higher wage. In that case, the excess of the wage paid a Visible with characteristics (e, η) over that paid to a similar Invisible is again nondecreasing in η : it is equal to zero for η between $\eta^*(e)$ and $\eta^{**}(e)$ and to $w(e, \eta) - w(e)$ outside that range. Figure 1 illustrates this case.

The observation that the difference between the wage of a Visible and an Invisible of equal ability η is increasing in η has important implications for the workers' incentives to acquire productive abilities of various kinds. We now turn to an analysis of the worker's choice of skill levels.

Suppose that a worker can accumulate human capital that enhances his or her ability but which is not directly observed by potential employers. Let ϵ denote the worker's expenditure on such human capital, and suppose that larger values of ϵ shift the distribution of η upwards (in the sense of stochastic dominance) in a way that is subject to initially large but eventually decreasing returns. More precisely, let $F(\eta|\epsilon)$ be the distribution of ability given ϵ , with $F(\eta_{\min}|\epsilon) = 1 - F(\eta_{\max}|\epsilon) = 0$ and $0 < F(\eta|\epsilon) < 1$ for $\eta \in (\eta_{\min}, \eta_{\max})$. We suppose that F has a corresponding density $f(\eta|\epsilon)$, that F is decreasing and strictly convex in ϵ , and that $\partial F(\eta|\epsilon)/\partial \epsilon$ evaluated at $\epsilon = 0$ is $-\infty$ for all $\eta \in (\eta_{\min}, \eta_{\max})$. Suppose, too, that workers invest in human

capital to maximize expected wages net of the expenditure ϵ . A formal analysis of the worker's problem (which is omitted) would rest on evaluations of first-order conditions. (The interested reader can easily supply the missing formal manipulations using the following fact: Our hypotheses about $F(\eta|\epsilon)$ imply that for any non-decreasing, function H that is not constant on $(\eta_{\min}, \eta_{\max})$, the integral $\int H(\eta) dF(\eta|\epsilon)$ is an increasing, concave, bounded function of ϵ , with an infinite derivative at zero.)

The first human capital choice situation we shall consider is when there is no η satisfying the condition (7) for promotion of an Invisible. Then, an Invisible's wage does not depend on ability, and it is clear that the optimizing Invisible will set $\epsilon = 0$. In the same situation, it is clear that Visibles will make positive unobservable investments in human capital, since the marginal return to a small investment is very large.

When there are some ability levels satisfying the promotion condition (7), the difference in (Visible minus Invisible) wages as a function of ability is, as noted earlier, a nondecreasing, non-constant function. So if Invisibles do not know their own abilities, the marginal returns to unobserved investments in human capital are higher for Visibles in that case, too. Next, suppose Invisibles do know their own native abilities (μ); for example let $\eta = \mu + \phi(\epsilon)$, where ϕ is an increasing function. Then the marginal return to ϵ is sometimes less, and never more, for Invisibles than for Visibles. Such reasoning leads to:

Proposition 3. Visibles invest at least as much in unobserved human capital as Invisibles with the same ability and education, regardless of whether workers know their own abilities at the time the investment is made. If workers do know their own abilities, then the ablest Invisibles always invest too little.

The foregoing analysis was carried out under the assumption that human capital is a one-dimensional variable. In the economy we have modelled, there are different kinds of jobs, and it therefore makes sense to assume that there are different kinds of human capital. In such a world, discrimination distorts not only the level of human capital investment, but also the kind of human capital acquired.

Consider the problem facing an Invisible, holding the observable education variable e fixed. Let us now regard the productivity parameters $\alpha(\epsilon_1, n)$ and $\beta(\epsilon_2, n)$ as functions of the unobserved choice $\epsilon = (\epsilon_1, \epsilon_2)$. The logic of the labor market equilibrium is unaffected by this change: It is still profitable for employers to hold back Invisibles with high productivity in low level jobs, because they are a source of rents, whether or not they have a comparative advantage in high level jobs.

Once again, there are two cases to be considered. In the first case, Invisibles with education level e are never assigned to high level jobs, and their pay in low level jobs does not depend on their unobservable productivity. In such a case, the question of investing in the wrong kind of human capital does not arise; Invisibles invest zero in both kinds.

The second possibility is that some Invisibles, but not the ablest ones, are promoted. In that case, if Invisibles do not know their own ability levels η , there is too little incentive for an Invisible to invest in the job skills ϵ_2 that give a comparative advantage in high level jobs. The reason for this is that Invisibles do not get the full benefit of the productivity enhancement in high level jobs, since with some probability they will be incorrectly assigned to low level jobs. Similarly, there is too little incentive to invest in generalized human capital ϵ_1 , because some of the productivity gain associated with that investment is squandered when the worker is assigned inefficiently to a low level job. If the two kinds of investments in human capital are complementary, then the effects noted in this paragraph are mutually reinforcing. Thus;

Proposition 4. Suppose that $\partial\alpha/\partial\epsilon_1$ and $\partial\beta/\partial\epsilon_2$ are positive, declining, and infinite at zero. Let $C(\epsilon_1, \epsilon_2)$ be the cost of acquiring unobservable human capital and suppose that it is increasing, differentiable and convex and that $\partial^2 C/\partial\epsilon_1\partial\epsilon_2 \leq 0$. Then Invisibles who make their human capital acquisition decisions without knowing their abilities set both ϵ_1 and ϵ_2 below their efficient levels.

The complementarity of costs assumed in Proposition 4 is not always realistic. For example, if a worker must decide how to allocate time between developing high level skills ϵ_2 or general skills ϵ_1 , and if costs are a convex function of the total time devoted, then the two activities will not satisfy the cost complementarity condition of Proposition 4. Compounding this

effect is the fact that acquiring skills which enhance productivity in low level jobs may reduce a worker's comparative efficiency in high level jobs. Typing skills provide one example. We may think of typing skills as increasing productivity in a low level job without affecting productivity in a higher level job, that is, increasing ϵ_1 and reducing the comparative efficiency parameter ϵ_2 . Then, acquiring typing skills makes promotion less probable than acquiring no skills at all. Invisibles, particularly the ablest ones, will find it in their interest to avoid acquiring such skills, or to hide them if they are already acquired.²

Although a strategy of avoiding skills that give one a comparative advantage in low level jobs may improve a worker's chances for promotion in the very imperfect world of our model, it is a deviation from the behavior required to achieve a fully efficient allocation. To the extent that these skills contribute not only to productivity in low level jobs, but to productivity in high level jobs as well, they will be acquired at inefficiently low levels. Moreover, despite this promotion-enhancing strategy, some of the able Invisibles will be assigned to low level jobs, and they will have acquired inefficiently low skill levels given their eventual assignments. Thus, the ablest Invisibles make distorted, inefficient choices; they acquire too little of the skills which weaken their comparative advantage in high level jobs.

It would be unreasonable to suppose that all human capital acquisition is unobserved, and it was for that reason that we included the observable education choice in our model of productivity. Unlike the choice of unobservable human capital ϵ , observable education e which raises a worker's productivity necessarily raises his or her wage. In general, in our framework, Visibles are paid their marginal products, while Invisibles are paid their expected maximal marginal products minus the efficiency loss that comes from the failure to promote the ablest of them. Thus, the marginal return to observable investments e in some kind of education will be higher for Invisibles than for Visibles if the more educated Invisibles suffer less discrimination (have less lost output due to inefficient assignments) than less educated ones. Conversely, if more educated Invisibles suffer more discrimination, then the marginal return is lower for Invisibles. To determine which of these cases applies, one must specify how investments in each kind of education affect productivity in the various kinds of jobs.

We can identify two theoretical cases in which the marginal return to observable investments education is lower for Invisibles. First, if the sole effect of the education in question were to enhance the productivity of the ablest workers in high level jobs, then since the ablest Invisibles are never assigned to high level jobs, the return to this sort of education for Invisibles would be zero. The second is the case where the equilibrium specifies that no Invisibles are ever promoted, and in which education enhances a worker's comparative advantage in high level work (so that $\partial\beta/\partial e > 0$). Then more educated workers

suffer more discrimination, that is, the efficiency loss from inappropriate assignments is greater for the better educated. In this case again, the marginal return to education is lower for the Invisibles than for Visibles.

Evidently, education is more valuable to an Invisible than to a Visible (i) if it enhances productivity in the low level job without enhancing productivity in the high level job, or (ii) if it is "remedial" in nature, raising the productivity of low-skilled workers in low level jobs, without equally raising the productivity of medium- and high-skilled workers in low level jobs. In both of these cases, education reduces the efficiency loss associated with improper job assignments at the market equilibrium.

It is interesting to interpret these findings in terms of the kinds of education that members of disadvantaged groups might seek to acquire. Invisibles who study a trade are likely to be more efficiently assigned, according to our theory, than those who acquire a general education: Potential employers can identify that an electrician assigned to mow lawns is underemployed, but cannot make the same determination for a clerk who would be more efficiently used as an office manager. Similarly, among those with college educations, those who study subjects like engineering acquire specific skills that make themselves easier to identify when they are underemployed. Such studies have higher returns for Invisibles, according to our theory, than a general education, even if the latter actually contributes more to the worker's productive capabilities.

PRIVATE AND PUBLIC STRATEGIES FOR REDUCING DISCRIMINATION

There is a class of workers who, as a result of their invisibility, are subject to wage discrimination. This discrimination brings with it a real, and potentially large, efficiency loss. The efficiency loss takes two forms. First, the most talented Invisibles will be inefficiently assigned to low level jobs. Secondly, Invisibles in general will acquire too little human capital and, moreover, they will acquire the wrong sorts of human capital.

Given that discrimination exists, we want next to consider whether there are viable strategies that either individual workers or government could use to reduce or eliminate this discrimination. We also want to consider whether there exist any private labor contracts which firms might make with members of disadvantaged groups which would both reduce discrimination and yield profits for the firm. Some economists argue that when such contracts exist, firms should be expected to offer them, creating a force that might work to undermine the discriminatory labor market equilibrium that we have described.

Perhaps the most obvious way for firms and workers to attempt to eliminate efficiency-distorting discrimination is for them to agree to a contract that binds the worker to the firm. If the firm were not threatened with a loss of rents from promoting a worker, it would have no motive to assign workers inefficiently. Then, the discriminatory mechanism we have identified would break down. It is therefore important to ask: To what degree can a worker be bound to a firm?

Even without special measures, there will be some natural bond between workers and firms emerging from the costs that workers incur by changing jobs, especially when that involves moving to a new community. Workers may also feel some degree of loyalty to a firm, especially just after a promotion. These bonds, along with the goodwill of employers, help to mitigate the effects we have described.

In addition, there are ways of strengthening the bond between a firm and its workers. Pension plans and deferred compensation plans of various kinds provide incentives for longevity in employment. Firms may develop a reputation for giving better, more flexible treatment to employees with long service. Salary schedules may be set to rise with an employee's term of employment, a device that effectively forces an employee to post a bond against quitting.

Creating strong bonds between a firm and its workers, however, also carries costs. Some job mobility is desirable, for the purpose of improved matching of workers to jobs or simply because of changing individual circumstances, such as marriage, the birth of a child, or a health problem. Also, an employer who has collected a bond, for example by paying low wages to new employees while paying wages to old employees above their marginal products, has an incentive to terminate the older employees or to make things so unpleasant for them that they will quit. These arguments are nearly identical to ones that are commonly raised in discussions of the inefficiently low employer investments in training employees, caused by the inability of

employers to retain trained employees without then immediately raising salaries to the market level.

It is fair to say, then, that some bonding of employees and firms is possible and desirable, and that this bonding will reduce the effects of the discrimination we have described by making it possible for employers to promote able Invisibles without losing all the rents earned from those workers. However, bonding is imperfect and incomplete. Certainly, there is no reason to presume that bonding is so good as to allow firms to eliminate discrimination against Invisibles.

Next, we turn our attention to how a government intervention might help to curtail discrimination. In an earlier version of this paper, we studied quota systems and wage scale interventions and found them to be ineffective against the kind of discrimination we have considered, assuming the government can only identify members of disadvantaged groups and not who is Invisible within the groups. The only potentially effective policy that we found — which combines promotion quotas with wage scales — is discussed below, together with a private market analogue to that policy. Throughout, we suppose (for the reasons discussed earlier) that there is a disproportionately large number of women and minorities among the Invisible group.

Consider a policy that combines quotas for the hiring and promotion of members of disadvantaged groups with a requirement that the average wage paid to these workers in high level jobs be the same as for the advantaged group. Such a policy would alter the incentives of individual firms in a complicated way. For example, it may pay a firm to hire a Visible woman executive at a

wage exceeding her productivity in order to bring up average wages for women managers while permitting continued discrimination against Invisible women in low level jobs. If the affirmative action program were narrowly applied, for example only to large government contractors, firms might achieve compliance by hiring and promoting able Visibles from the disadvantaged groups. Such a tactic would leave the wage paid to each worker unchanged and would carry no benefit for any disadvantaged workers. If, however, the government program were applied broadly, it would be impossible for all firms to adopt such "neutral" tactics, so the labor market equilibrium would be upset.

How might an effective quota-and-wage-scale policy be devised? For any fixed distribution of abilities among Invisibles with education e , there is some assignment of workers associated with efficient production. Let $p^*(e)$ be the efficient proportion to be promoted and let $w^*(e)$ be the average marginal product associated with the efficiently promoted group. Consider a policy that mandates the assignment of the proportion $p^*(e)$ of disadvantaged workers to high level jobs with an average wage of at least $w^*(e)$. Then an individual firm's wage bill cannot be reduced by assigning workers to jobs inefficiently, so a profit-maximizing firm would make efficient assignments. In particular, it is an equilibrium for firms to assign all workers to jobs efficiently, to pay Visibles and promoted workers their marginal products, and to pay unpromoted Invisibles their expected marginal products in the job they occupy.

To implement such a policy, the government would have to know the efficient proportion to promote (p^*) and their average marginal product (w^*) when all assignments are made efficiently. If, the government did not know these magnitudes, it would have to base its policy on guesses about them. For example, it might assume that average productivities are the same among women and minorities as among, say, white males. Of course, any such assumption would inevitably be wrong, and inefficiencies would be the result.

Now suppose that the program does set a promotion quota and wage scale designed to raise the average status and pay of women and minorities to the same level as for other groups. Suppose, too, that the labor market moves to a new equilibrium. This equilibrium still provides too little incentive for Invisibles to acquire unobservable human capital, since the wage paid to those in low level jobs does not vary with ability. Even if the government were omniscient in applying this program, full efficiency could not be reached. Nevertheless, the quota-and-wage-scale program could increase output per firm and reduce differences in average wages and in promotion probabilities.

The dynamics³ that result from introducing an efficient quota-and-wage-scale program are particularly interesting. When first introduced, given the pre-existing promotion policy and resulting patterns of human capital acquisition, the average skill level of the female and minority managers would be less than the equilibrium level, so there would be too few qualified Invisibles to fill the high level jobs. Resistance and resentment could develop based on the accurate perception that the

initial beneficiaries of the affirmative action policy were unqualified and overpaid relative to their coworkers. In the long run, however, if the quotas and wage scales were set properly, the new pay and promotion policies would lead to more similar patterns of human capital acquisition across groups, so that the actual productivity of disadvantaged workers would rise. Thus, in the long run, wages and job assignments would end up being more nearly in line with ability level.

Part of the interest in the foregoing government program is to see whether there are any contracts possible in the private sector that achieve essentially the same equilibrium results. Thus, suppose firms offer employment contracts that specify a fixed probability of promotion and an average wage conditional on promotion. Suppose, too, that the initial distribution of abilities were the same for Visibles and Invisibles. One equilibrium would be for all firms to offer Invisibles a probability of promotion $p^*(e)$ and to promise an average wage of $w^*(e)$ to those promoted. The corresponding assignments and the actual wages paid would be much the same as the one predicted to result from an omniscient government-imposed quota-and-wage-scale policy.

Contracts such as we have described do not exist in the private labor market. In part this may be a result of private enforcement costs, since the contract depends on information concerning the wages paid to individuals within the firm, which the firm may be reluctant to reveal. A more fundamental reason is that contracts of this sort cannot overturn the simpler, but less efficient equilibrium we have studied above. If a single

firm offered such a contract, it could meet its obligation to promote, say, women or blacks and pay them well by hiring the ablest Visible women or blacks at the going wage, without doing anything to improve the lot of its Invisibles. Its incentives to discriminate against Invisibles are unchanged, and so, at equilibrium, it would have no advantage in attracting workers. A firm offering such a contract would not upset the status quo. The key observation here is that *the desired contracts enhance efficiency only if many firms offer them*. This is analogous to the observation that government programs of limited applicability fail to provide incentives to eliminate discrimination. There is no market force that would lead all firms to offer such contracts simultaneously, upsetting the equilibrium.

The government program required to achieve efficiency is quite obtrusive, and requires knowledgeable sensitive implementation to be effective. Let us therefore turn to an analysis of a policy directed more narrowly at the inefficient acquisition of skills by Invisibles, to see how much can be done on that dimension. A plausible policy instrument might be an education subsidy, available to members of the disadvantaged groups who make up a disproportionate part of the Invisible population.

Little can be done to improve the direct incentives to acquire unobservable human capital. Since these acquisition activities are unobserved, it is impossible to base a subsidy directly on them. One can, however, subsidize education and other *observable* investments in human capital. The direct effects of subsidies to education will be either enhanced or reduced depending on whether unobservable investments in human

capital are complements or substitutes for the observable investments.

To put the case for educational subsidies in its most favorable light, let us assume that unobservable human capital investments are of negligible importance (so that unobservable choices are not distorted), that workers do not know their own abilities (so that distortions in the kind of education chosen within the Invisible group are unimportant), that groups can be identified that consist almost solely of Invisibles (so that the subsidies can be well targeted), and that the subsidies are set at a level that leads Invisibles to make the same education choices as Visibles. Then, the distribution of education and skills will be the same for the Visibles and Invisibles. But, Proposition 2 still applies: Invisibles will still earn less and be promoted less often than Visibles. Education subsidies may improve the lot of Invisibles, but they can do nothing to eliminate or even ameliorate the basic problem of discrimination by employers. Moreover, given our assumptions and given the discrimination in job assignments, the education choices made by unsubsidized workers is actually optimal. So education subsidies actually increase the inefficiencies associated with discrimination in job assignments.

Finally, we consider whether there are any strategies which individuals in the disadvantaged group could use to reduce the discrimination they experience. One possible answer is for the disadvantaged to try to gain visibility. Indeed this has been a thrust of both the civil rights movements and the womens' move-

ment, as both groups have sought to reduce social discrimination, arguing strongly that such discrimination is one of the causes of job and wage discrimination.⁴

A second possibility is in the area of education. It may be that some kinds of education serve directly to enhance visibility. One might then expect members of disadvantaged groups to choose those kinds of education. This may help explain, for example, why minorities receive a disproportionate number of advanced degrees in engineering.

In the labor-market signalling model of Spence (1973), education serves as a signal of a worker's ability, and therefore makes the worker visible. Spence's argument is based on the assumption that workers have private information about their abilities, and that, *ceterus paribus*, abler workers would choose more education than less able workers. In our model, however, discrimination can take place even when workers have no private information about their abilities at the time of taking an education decision. The effectiveness of signalling depends on assumptions about the labor market which are otherwise irrelevant to the effects studied in our model.

CONCLUSION

We have shown how non-economic forces arising from segregation and misperceptions can impinge on the functioning of a labor market, resulting in discrimination by employers against employees whose backgrounds, personalities, and connections cause their skills to be largely unnoticed by potential employers. According to our theory, employers earn rents on the most talented individual Invisibles, but they do not earn rents on Invisibles in total. All losses suffered by Invisibles are "dead-weight" losses associated with improper job assignments and inappropriate human capital acquisition decisions. Unlike Becker's theory, which holds that discrimination is efficient, ours holds that discrimination causes a loss to society as a whole rather than just a transfer from Invisibles to Visibles.

The conclusion that the losses suffered by Invisibles are dead-weight losses has important implications for the politics of implementing affirmative action programs. In the long run, well-designed programs will have a neutral effect on firms, since firms have no economic interest in perpetuating discrimination. In the short run, however, firms may be made worse off by aggressive affirmative action policies, at least until Invisible workers can acquire the skills needed for their new assignments. Thus firms are likely to be most opposed to short time-table programs, but may be neutral toward more gradual approaches. Similarly, since only economy-wide changes can be effective for eliminating discrimination whereas action directed at individual firms can impose losses on them, firms can be expected to resist strongly programs which are narrowly directed.

Two of the predictions of our theory are that wages are more variable among Invisibles than among Visibles and that Visibles are more able to increase their wages by investing in human capital. These predictions, if corroborated, may be as important as the general lower level of wages for understanding the effects of economic discrimination on the noneconomic aspects of life in disadvantaged groups. For example, the greater wage dispersion among men than women increases the relative importance of earning capacity as a characteristic of men for the purpose of attracting a spouse. The man's greater ability to affect his personal earning capacity by his own efforts further contributes to the emphasis among men on finding a well-paying job or career. For women, if differences in earning power are indeed less marked, they will be less important for attracting a suitable mate. Competition among women for spouses must then be directed along other dimensions than earning power. Among minorities, where the male wages are predicted to be lower, less variable, and less easily raised by diligence than for more advantaged groups, men, too, will be more likely to compete for spouses along dimensions other than earning power.

Our explanation of the economics of job discrimination is not intended to be self-contained. Becker's "taste for discrimination" is an important factor in societies where prejudice is deep and widespread and the disadvantaged group is large. But what of the cases where prejudice is weaker and far from universal? Why are market forces then seemingly unable to ensure that women, minorities, and other disadvantaged workers are efficient-

ly employed and paid their maximal marginal product? We have found that market forces are not only weakened by the Invisibility Problem, they are actually twisted to give employers a profit motive for inefficient discrimination against Invisibles.

REFERENCES

- Kenneth J. Arrow, "Models of Job Discrimination," in A.H. Pascal (ed), *Racial Discrimination in Economic Life*, Lexington: Lexington Books, 1972.
- Gary S. Becker, *The Economics of Discrimination*, Chicago: University of Chicago Press, 1957.
- Rosabeth M. Kanter, *Men and Women of the Corporation*, New York: Basic Books, Inc., 1977 (especially page 99).
- Shelly J. Lundberg and Richard Startz, "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review*, Vol. 73, 1983, 340-347.
- Paul R. Milgrom and Robert J. Weber, "A Theory of Auctions and Competitive Bidding," *Econometrica*, Vol. 50, 1982, 1089-1122.
- Edmund S. Phelps, "The Statistical Theory of Racism and Sexism," *American Economic Review*, Vol. 62, 1972, 659-661.
- A. Michael Spence, "Job Market Signalling," *Quarterly Journal of Economics*, Vol. 87, 1973, 355-374.
- Michael Waldman, "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics*, Vol. 15, 1984, 255-267.

END NOTES

¹The assumption of no cuts and no layoffs can be derived from a more detailed model in which there are large firm specific training costs. Our assumption that contracts simply specify a wage, rather than, say, a wage and a promise that a certain proportion of women and minorities will be promoted, is more restrictive. Such promises might be made directly, in the employment contract, or indirectly, through a reputation that the firm acquires from its actual behavior toward employees. The consequences of such promises will be explored later in the paper.

²This corresponds well with the common wisdom among female managers that it is unwise to acquire good typing skills, because good typists are inevitably assigned to do the typing.

³These are market dynamics in the old-fashioned sense of analyzing short-run and long-run equilibrium. Our model is purely static.

⁴For example, concerns about visibility are a fundamental part of the public policy issue in the lawsuit against the all-male Jaycees business club.

APPENDIX
Proof of Proposition 1

We need to show that the only wage functions and production policies consistent with equilibrium are those given by the Proposition, and that the given wages and policies are an equilibrium. To do this, we first break the firm's optimization problem into two parts: (i) wage cost minimization given the levels x and y of high level and low level labor to be used and (ii) profit maximization by choice of x and y using the cost function resulting from part (i).

Let $W(x,y)$ denote the minimum total wage cost of acquiring x units of low level labor and y units of high level labor, where the firm's instruments are its hiring policy functions m and n , and its promotion policy functions p and q . Fix the numbers x and y , both strictly positive. We perform a change of variables to work directly with the numbers (rather than proportions) of each kind of worker assigned to each kind of job. Thus, we formulate the mathematical problem as one of minimizing wage costs by choice of the variables $m(e)$ and

$$a(e,\eta) = n(e,\eta)[1-p(e,\eta)]$$

$$b(e,\eta) = n(e,\eta)p(e,\eta)$$

$$c(e,\eta) = m(e)[1-q(e,\eta)]h(\eta|e)$$

The mathematical statement of the problem is to minimize the wages:

$$(A0) \quad \iint [(a(e,\eta) + b(e,\eta)) w(e,\eta) + c(e,\eta) w(e)] d\eta de \\ + \iint (m(e)h(\eta|e) - c(e,\eta)) \max[w(e), w(e,\eta)] d\eta de$$

subject to the constraints

$$(A1) \quad \iint \alpha(e,\eta) [a(e,\eta) + c(e,\eta)] d\eta de \geq x$$

$$(A2) \quad \iint \alpha(e,\eta) \beta(e,\eta) \{b(e,\eta) + m(e) h(\eta|e) - c(e,\eta)\} d\eta de \geq y$$

$$(A3) \quad m(e) h(\eta|e) - c(e,\eta) \geq 0 \quad \text{for all } (e,\eta)$$

$$(A4) \quad a, b, c, m \geq 0$$

As now formulated, the cost-minimization problem has a linear objective and linear inequality constraints; it is a continuous linear program. To prove the optimality of a solution of the specified form, it therefore suffices to identify dual variables (Lagrange multipliers) λ_x , λ_y , $\mu(e,\eta)$ for the constraints (except the non-negativity constraints) such that program is primal and dual feasible and the complementary slackness conditions of linear programming hold. "Dual feasibility" means that the Lagrange multipliers must satisfy constraints (A5)-(A9) below. In what follows, we denote the wage $\max[w(e), w(e,\eta)]$ paid to a high level Invisible by $M(e,\eta)$.

$$(A5) \quad w(e,\eta) - \lambda_x \alpha(e,\eta) \geq 0 \quad \text{for all } (e,\eta)$$

$$(A6) \quad w(e,\eta) - \lambda_y \alpha(e,\eta) \beta(e,\eta) \geq 0 \quad \text{for all } (e,\eta)$$

$$(A7) \quad w(e) - M(e,\eta) - \lambda_x \alpha(e,\eta) + \lambda_y \alpha(e,\eta) \beta(e,\eta) + \mu(e,\eta) \geq 0 \\ \text{for all } (e,\eta)$$

$$(A8) \quad \int [M(e,\eta) - \lambda_y \alpha(e,\eta) \beta(e,\eta) - \mu(e,\eta)] h(\eta|e) d\eta \geq 0 \\ \text{for all } e$$

$$(A9) \quad \lambda_x \geq 0, \lambda_y \geq 0, \mu(e,\eta) \geq 0 \quad \text{for all } (e,\eta)$$

"Complementary slackness" means that λ_x , λ_y , or $u(e, \eta)$ can be positive only if (A1), (A2), or (A3) respectively holds with equality, and similarly $a(e, \eta)$, $b(e, \eta)$, $c(e, \eta)$, or $m(e)$ can be positive only if (A5), (A6), (A7), or (A9) holds with equality.

Market clearing requires that either $a(e, \eta) > 0$ or $b(e, \eta) > 0$, so by complementary slackness either (A5) or (A6) holds as an equality. It follows immediately that

$$w(e, \eta) = \max(\lambda_x \alpha(e, \eta), \lambda_y \alpha(e, \eta) \beta(e, \eta)),$$

which proves that the Visible wage function takes the form given by equation (4) with $f_1 = \lambda_x$ and $f_2 = \lambda_y$.

Next we establish the optimality of the promotion policy for Visibles. Suppose some (η, e) -Visibles are hired by not promoted at the optimum. Then $a(e, \eta) > 0$, so by complementary slackness (A5) holds with equality. Given the form of the wage function established in the preceding paragraph, this proves that

$$(A10) \quad \lambda_x \alpha(e, \eta) < \lambda_y \alpha(e, \eta) \beta(e, \eta).$$

Since β is increasing in η , using the definition (9) of $\eta^*(e)$, (A10) is never satisfied when $\eta > \eta^*(e)$, so *all such Visibles are promoted*. Next, suppose some (η, e) -Visibles are promoted at the optimum. Then $\beta(e, \eta) > 0$, so (A6) holds with equality, and the inequality (A10) is reversed. Since β is increasing in η , the reverse inequality is never satisfied when $\eta < \eta^*(e)$, so *no such Visibles are promoted*. Together, the two italicized observations are equivalent to the promotion policy (6).

For Invisibles, the number of unpromoted (η, e) 's is the left-hand side of (A3). If any are promoted, then (A3) holds as a strict inequality, so $\mu(e, \eta) = 0$ by complementary slackness. Then, using (A7), we have:

$$(A11) \quad w(e) - M(e, \eta) - \lambda_x \alpha(e, \eta) + \lambda_y \alpha(e, \eta) \beta(e, \eta) \geq 0.$$

This cannot hold for $\eta < \eta^*(e)$ since, for that case, $M(e, \eta) \geq w(e)$ and $\lambda_x \alpha > \lambda_y \alpha \beta$. Similarly, it cannot hold for $\eta > \eta^{**}(e) > \eta^*(e)$. Thus, *no such Invisibles can be promoted at the optimum*. For the converse, suppose some Invisible (η, e) 's are unpromoted at the optimum. Then $c(e, \eta) > 0$, and hence (A7) holds with equality. Since $\mu(e, \eta) \geq 0$, this means that (A11) must hold with the inequality reversed. However, given the definitions (9) and (10) of $\eta^*(e)$ and $\eta^{**}(e)$, this cannot occur when $\eta^*(e) < \eta < \eta^{**}(e)$, so *all such Invisibles are promoted at the optimum*. Together, the two italicized observations are equivalent to the promotion policy (7).

Market clearing for the Invisibles implies that $m(e)$ is positive and so (A8) holds with equality. This leads, after some algebra, to the Invisibles' wages (5).

From the form of the solution given above, we see that $W(x, y) = f_1 x + f_2 y$. (The dual prices of (A1) and (A2) are f_1 and f_2 independent of x and y .) Hence, the firm's problem in choosing x and y , which is to maximize $f(x, y) - W(x, y)$, leads to the first order necessary conditions (8).

Thus, we have shown that (4)-(8) are necessary for equilibrium. The proof of sufficiency is easier. To show that (4)-(8) together with market clearing are sufficient conditions, it suffices to show that the proposed solution to the firm's problem is indeed an optimal one. To show that the wage cost minimization solution is optimal, it suffices to display values for the dual variables and verify that the primal constraints, the dual constraints, and complementary slackness are all satisfied. The values for the dual variables are given by (A11)-(A13).

$$(A11) \quad \lambda_x = f_1$$

$$(A12) \quad \lambda_y = f_2$$

$$(A13) \quad \mu(e, n) = \max[0, M(e, n) + f_1 \alpha(e, n) - f_2 \alpha(e, n) \beta(e, n) - w(e)]$$

We omit the verifications. The only difficult one is (A8), which (using (4)) can be shown to hold with equality for all e .

As noted above, $W(x, y) = f_1 x + f_2 y$. Hence the firm's problem, stated below, is concave:

$$\begin{array}{l} \text{Max } f(x, y) - W(x, y) \\ x, y \end{array}$$

The need for positive amounts of labor in both kinds of job ($x > 0$, $y > 0$) for production ensures that, at the market clearing solution, the firm is at an interior optimum. The first-order optimality conditions (8) are therefore sufficient as well as necessary.