Diversity in the Workplace

By John Morgan and Felix Várdy*

An unbiased employer engages in optimal sequential search by drawing from two equally qualified subpopulations of job candidates who differ in their "discourse systems." That is, minorities convey noisier unbiased signals of ability than non-minorities. We show that when the employer is selective, minorities are underrepresented in the workforce, fired at greater rates, and underrepresented among initial hires. Workplace diversity increases if: the cost of firing falls, the cost of interviewing increases, the opportunity cost of not hiring increases, or the average skill of candidates increases. If, however, the employer is sufficiently unselective, minorities may be *overrepresented* in the workforce.

JEL Numbers: D21, D63, D83, J71, J78

Keywords: Diversity, sequential search, statistical discrimination, reverse discrimination, discourse systems, integration of minorities.

A central social, political, and economic challenge confronting the European Union today arises from the tensions created by the growth of immigrant populations, particularly those from predominantly Muslim countries. These tensions have manifested themselves in sometimes dramatic fashion—the murder of Theo van Gogh and its aftermath in the Netherlands, the widespread unrest over Danish anti-Muslim cartoons, and the weeks-long violence and rioting in the outskirts of Paris in the Fall of 2005. Many have argued that these events are mere symptoms of a broad underlying discontent caused, in large part, by a lack of economic opportunities. Indeed, generally high unemployment in the European Union (EU), often attributed to labor market rigidities, affects immigrant populations particularly severely: unemployment rates for minorities remain stubbornly higher than for the majority, and grow especially severe during economic downturns.

What accounts for the disparity in the employment experiences of Europe's majority populations versus its minority populations? Skill and age differences are surely part of the explanation. Minority populations are, on average, less educated and younger than the majority, and unemployment rates tend to be higher among the low skilled and the young. Nevertheless, while the employment disadvantage of minorities is reduced once differences in educational attainment and age are taken into account, it does not disappear. (See, e.g., Paul Tesser, Ans Merens and Carlo van Praag, 1999, and Jaco Dagevos, 2006). Of course, it may be that employers simply have a taste for discrimination and that the underrepresentation of Muslims in the EU workforce reflects the strength of these tastes. While it is hard to rule out this explanation, one would expect to see the unemployment gap between Muslims and the rest of the population shrink as competitive pressures from outsourcing and globalization have increased. In fact, the opposite has occurred in the Netherlands over the last couple of years (Dagevos, 2006).

An alternative explanation for higher minority unemployment may be gleaned from the sociolinguistics literature. (See, for instance, Ronald Scollon and Suzanne Scollon, 2001.) According to this hypothesis, minority job candidates struggle to make themselves understood due to differences in "discourse systems." For example, a candidate's behavior during a job interview may be quite revealing to an employer if they share the same social or cultural background. But if they do not, it can be much harder for the employer to form an accurate opinion about the applicant. In other words, the signals conveyed by minorities during interviews may be so garbled that they fail to convince (majority) employers of their qualities, even when they are perfectly competent and employers have no taste for discrimination. In contrast, by virtue of sharing the same discourse system as employers, majority job candidates do not face this problem. Thus, for them it tends to be easier to convey an accurate impression of their ability. As a consequence, minority populations find greater difficulty in securing employment than majority populations.[1]

This hypothesis raises several questions: can differences in discourse systems alone explain differences in unemployment rates between majority and minority populations, absent any differences in underlying ability of the two populations? If so, what policy prescriptions

[1] Of course, matching the background of the interviewer with the background of the candidate would solve this problem. However, more often than not, this may be quite difficult to implement. First, in organizations lacking diversity, minorities are scarce to begin with. Second, it should not be forgotten that the various minorities are culturally highly diverse, thus requiring a very careful matching between the evaluator and the evaluee. For instance, while a French speaking West African and an African-American are both people of color, it seems quite clear that they do not share the same discourse system.

could remedy this? Should employment protection be increased or decreased? What about other rigidities—are these helpful or harmful to workplace diversity? What about macro implications—can the EU simply grow itself out of the problem?

To examine these questions, we study a model in which an employer tries to fill a vacancy by sequentially interviewing job candidates from a pool of potential employees. The pool consists of two subpopulations. One subpopulation may be thought of as the majority population, the other as the minority population. The employer has no inherent taste for discrimination and the only thing he cares about is whether a candidate can do the job. On average, candidates from both subpopulations are equally likely to be able to do the job. Candidates do, however, differ in their discourse systems. To capture this difference, we suppose that when the employer interviews a minority candidate he receives a noisier signal of that candidate's true ability than when he interviews a majority candidate.

Our main result shows that, when an employer is "selective," equilibrium always entails underrepresentation of the minority population in the permanent workforce. Here, "selective" means that only candidates are hired for whom the post-interview probability that they can do the job exceeds the prior probability. More surprisingly, when an employer is sufficiently "unselective," equilibrium entails *over*representation of the minority population. Sufficiently "unselective" means that a candidate is hired provided he does not disappoint too much during the interview. Finally, regardless of the selectivity of the employer, the firing rate of minority candidates always exceeds that of majority candidates.

The intuition for the main result may be seen in the following example. Suppose that the prior probability that a random candidate can do the job is 50% and assume that the employer is *selective*, such that only candidates about whom the employer is at least 95% certain after the interview that they can do the job are hired. (Such a high threshold may be optimal when firing costs are very high.) In that case, the relative uninformativeness of a minority candidate's signal about his qualifications makes it extremely hard to change the employer's 50% prior belief of "success" to a posterior belief of at least 95%. Therefore, it is very unlikely that a minority candidate is going to fill the position. As a result, selective hiring practices lead to severe underrepresentation of minorities, even though minorities are as competent as the majority and employers are not prejudiced against them. On the other hand, if the employer is not selective at all, such that any candidate is hired provided that the posterior probability that he can do the job is no less than 5%, then the relative uninformativeness of a minority candidate's signal about his qualifications is an advantage. It makes it virtually impossible for the employer's 50% prior belief of success to be downgraded to less than 5%. Under these circumstances, virtually all minority candidates are given a chance and remain in the job if they turn out to be good. At the same time, in relative terms, many majority candidates are turned away at the gate, because the informativeness of their signals does make significant belief revisions possible. As we show, this leads to "reverse discrimination:" minorities will be overrepresented in the workforces of unselective employers. For similar reasons, the model also predicts that the degree of underrepresentation of minorities depends on the prior probability that random candidates can do the job. Specifically, minorities will be most severely underrepresented in positions that demand rare skills, such that the employer's priors are very pessimistic. In contrast, minorities will be overrepresented in positions that nearly anyone can do.

Next, the model predicts that the relative representation of minorities in the workplace varies over the business cycle. Specifically, if employers are at all selective, diversity is predicted to be *procyclical*, increasing during economic upturns and decreasing during downturns. Intuitively, when the economy is booming, recruiting job candidates is more costly. At the same time, the opportunity cost of leaving the position unfilled is higher. Both effects make the employer less picky, encouraging employers to "take a chance" on job candidates whose quality is uncertain. This reduces the underrepresentation of minorities.

This prediction is roughly consistent with the Dutch experience over the last decade. During the second half of the 1990s, a period a rapid economic expansion, unemployment among Muslim minorities in the Netherlands fell quite spectacularly, from over 30% in 1995 to around 9% in 2001. During the same period, the unemployment rate among the non-immigrant Dutch fell from around 6.5% to 3%. Since then, the trend has largely reversed. By 2005, unemployment among Muslims was again as high as 24%, while unemployment among the non-immigrant Dutch had only risen to 5%. (Dagevos, 2006.)

Finally, we turn to policy solutions to the "diversity problem." Our main finding in this regard is that high firing costs harm diversity. Intuitively, protections that raise the cost of firing lead the employer to guard more vigilantly against Type II errors (hiring of incompetent candidates). The employer achieves this by becoming more selective, which exacerbates the underrepresentation of minorities. This suggests that labor market rigidities such as high costs of firing contribute to the economic and social exclusion of Muslim minorities in Europe.

To conclude, the model implies that differences in discourse systems can indeed generate differences in unemployment across otherwise homogeneous populations. Going beyond the model, it suggests a feedback system between cultural and economic barriers to integration: the lack of a shared discourse system leads to few opportunities for minorities to land demanding jobs with selective employers. Instead, minorities are more likely to be unemployed, or stuck at the lower end of the labor market. This, in turn, implies that they are less likely to be in close contact with the dominant discourse system and, therefore, the cultural segregation across populations is self-reinforcing and does not disappear over time.

Related Literature

The nearest antecedent to our paper is Bradford Cornell and Ivo Welch (1996), who look at minority hiring when the employer uses a fixed-sample search strategy. As is standard in the fixed-sample search literature (see, for instance, George Stigler, 1961), order statistics determine the main economic effects. In contrast, we employ an optimal sequential search approach in the spirit of John J. McCall (1970). This allows us to explicitly model and analyze the effects of what Cornell and Welch call "ex ante screening" versus "on-the-job performance measurement".

Our work is also related to the statistical discrimination literature starting with Edmund S. Phelps (1972). Unlike our model, this literature assumes that majority and minority populations differ statistically with respect to some *payoff relevant* characteristic, such as average labor productivity. While Phelps studied models where the population means of the payoff relevant characteristic differ, his work has been extended to study differences in variances as well (Dennis Aigner and Glen Cain, 1977). More recently, by endogenizing human capital acquisition, Stephen A. Coate and Glenn Loury (1993) as well as Shelly J. Lundberg and Richard Startz (1998) have shown how statistical discrimination can arise even with ex ante homogeneous populations.

Less closely related to our work is the discrimination literature that assumes that employers inherently dislike minorities. (See, e.g., Gary S. Becker (1957), as well as Kenneth J. Arrow (1998) for a survey.) Somewhat related to our work is Dan Black (1995), who examines this motive in a search-theoretic setting, and Asa Rosen (1997) who combines search with a match-specific payoff. Finally, we should mention other "language theories of discrimination," such as Kevin Lang (1986) or Susan Athey, Christopher Avery, and Peter Zemsky (2000). These models rely on communication complementarities within a firm as opposed to our model, where the focus is on communication between an employer and job candidates.

I. Model

We study a labor market search problem in which the employer does the searching. In order to fill a vacancy, an employer takes random draws at a cost $k > 0$ per draw from a countably infinite population of job candidates. Each draw can be thought of as the employer conducting a job interview with a candidate. Each candidate has two characteristics: what subpopulation he belongs to, which is observable to the employer at the time of the interview; and whether he can do the job, which only becomes observable if the candidate is actually hired. We shall refer to the former characteristic as a candidate's *kind* and to the latter as a candidate's *type*.

A candidate's kind is denoted by $\gamma \in \{A, B\}$. A fraction m_A of the candidates are from subpopulation A, which consists of members of the "dominant" culture—i.e., candidates with the same discourse system as the employer/evaluator. The remaining fraction $m_B = 1 - m_A$ of the candidates are from subpopulation B, which consists of members *not* belonging to the dominant culture. As a shorthand for differences between the dominant and non-dominant cultures, we shall sometimes refer to candidates of kind A as "majority" candidates and candidates of kind B as "minority" candidates although, as the description above makes clear, majority candidates do not necessarily have to be more numerous than minority candidates.

A candidate's type, denoted by θ, equals 1 if he can do the job and equals zero if he cannot. Let p_γ denote the probability that a randomly drawn candidate of kind γ can do the job; that is, $p_\gamma \equiv \Pr(\Theta = 1|\gamma)$. We assume that the two subpopulations are equally qualified to do the job, that is, $p_A = p_B = p$. Hence, none of the results in the paper are driven by differences between the type distributions in the subpopulations.

In advance of the interview, the employer does not know or does not act upon information as to whether a candidate is a member of the dominant culture or not.[2] However, at the interview stage, a candidate's kind —A or B— is perfectly revealed to the employer through some easily observable characteristic such as dialect or skin color. In addition, the interview also reveals to the employer a signal S_γ as to the competence of the candidate, where $S_\gamma = \theta + \varepsilon_\gamma$. That is, the signal is equal to the candidate's type θ plus an error term ε_γ, which is assumed to be Normally distributed with zero mean and variance σ_γ^2.

The key difference between candidates of different kinds is that the employer finds it easier to assess the competence of candidates from the same culture as compared to those from a different culture. To model this difference, we assume that $\sigma_B > \sigma_A$. That is, from the perspective of the employer, there is more noise in the signal of a minority candidate than in the signal of a majority candidate.

The timing of the employer's decision problem is as follows. In period 1, the employer draws a random candidate and conducts an interview at a total cost k. On the basis of the candidate's interview signal s, and taking into account his kind γ, the employer calculates the candidate's "success probability" q. That is, q is the employer's posterior belief about the probability that the candidate can do the job. Given q, the employer then decides whether to hire the candidate and period 1 ends.

In period 2 and all subsequent periods, if the employer did not hire in the previous period he interviews a new candidate and the process proceeds as before. If, however, the employer did hire in the previous period, the employee's type θ is perfectly revealed to the employer. If the employee can do the job, i.e., $\theta = 1$, he is retained forever and all search ceases. In that case, the employer enjoys a payoff with a net present value of $v > 0$. If, however, the

[2] In reality, an employer may be able to guess a potential candidate's minority status from his name or address. On the basis of that information, the employer might decide not to invite him for an interview. Even though in most countries this is clearly against the law, there is evidence that it does happen. See, for example, Bertrand and Mullainathan (2004). The assumption in our model is that employers do abide by the law and, therefore, do not discriminate in this way. Technically speaking, our model is one of undirected search.

employee cannot do the job, i.e., $\theta = 0$, then by retaining the employee the employer earns a payoff with a net present value of $-w < 0$. Alternatively, the employer can fire the employee in period 2 and incur a cost of $c > 0$. Throughout, we assume that $c < w$, such that it is always optimal to fire incompetent employees. Finally, we assume that the employer has a discount factor $\delta \in (0, 1)$ between periods.

Posterior Beliefs

As we shall see, the employer's optimal strategy is to impose a uniform success probability threshold, $\underline{q}^*$, when deciding whether to hire a candidate. That is, a candidate is hired if and only if the probability that he can do the job is at least $\underline{q}^*$. The optimal threshold depends on the posterior distribution of the employer's beliefs as to the competence of a candidate. Thus, it is useful to summarize key features of this posterior distribution.

Define $q_\gamma(s)$ to be the employer's posterior belief that a candidate of kind γ with signal s can do the job; that is, $q_\gamma(s) \equiv \Pr(\Theta = 1 | S_\gamma = s)$. By Bayes' rule, we can rewrite this expression as

$$q_\gamma(s) = \frac{\phi\left(\frac{s-1}{\sigma_\gamma}\right) p}{\phi\left(\frac{s-1}{\sigma_\gamma}\right) p + \phi\left(\frac{s}{\sigma_\gamma}\right)(1-p)}$$

where $\phi(\cdot)$ denotes the density of a standard Normal random variable.

It will sometimes be useful to determine the signal realization s corresponding to a given success probability q, which we shall denote by $s_\gamma(q)$. Since $q_\gamma(s)$ is a monotone function, it is invertible in the extended reals and $s_\gamma(q)$ is well-defined. Using that $\phi(t) \equiv \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}t^2\right]$, it may be readily shown that

$$s_\gamma(q) = \frac{1}{2} - \sigma_\gamma^2 \ln\left(\frac{1-q}{q}\frac{p}{1-p}\right)$$

Prior to the realization of the signal but after having observed a candidate's kind, the success probability $Q_\gamma = q_\gamma(S_\gamma)$ is a random variable. Now, let $G_\gamma(\cdot)$ denote the cumulative distribution function (cdf) of Q_γ. Formally,

$$G_\gamma(q) = p\,\Phi\left(\frac{s_\gamma(q) - 1}{\sigma_\gamma}\right) + (1-p)\,\Phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right)$$

where $\Phi(\cdot)$ denotes the cdf of a standard Normal distribution. The associated density of $G_\gamma(q)$ is

$$g_\gamma(q) = \left(p\phi\left(\frac{s_\gamma(q) - 1}{\sigma_\gamma}\right) + (1-p)\,\phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right)\right)\frac{\sigma_\gamma}{q(1-q)}$$

Similarly, let $G(\cdot)$ denote the cdf of success probability Q prior to observing the candidate's kind or signal, and $g(\cdot)$ denote the associated density. Formally, $G(q) = (1 - m_B) G_A(q) + m_B G_B(q)$. Finally, it is useful to establish the following stochastic dominance relations for $G(\cdot)$ and $G_\gamma(\cdot)$.

Lemma 1 *For all $p > p'$, $G(\cdot\,; p)$ first-order stochastically dominates $G(\cdot\,; p')$. That is, $\frac{d}{dp} G(q) < 0$, for all $q \in (0, 1)$.*

Lemma 2 *$G_A(\cdot)$ is a mean preserving spread of $G_B(\cdot)$. And, for all $m_B < m'_B$, $G(\cdot\,; m_B)$ is a mean preserving spread of $G(\cdot\,; m'_B)$.*

II. Optimal Search and Hiring

In this section, we show that there exists a unique solution to the employer's optimization problem. The optimal hiring strategy is to set an identical success probability threshold, $\underline{q}^*$, for all candidates irrespective of their kind. That is, after observing a signal s from a candidate of kind γ, the candidate is hired if and only if the posterior probability that he can do the job, $q_\gamma(s)$, is at least $\underline{q}^*$.

To see this, let V^* denote the employer's expected payoff if he follows an optimal search and hiring strategy. In any optimal strategy, the employer hires a candidate if and only if his belief q that the candidate can do the job is such that the expected payoff from hiring, which we denote by $H(q, V^*)$, exceeds the expected payoff from not hiring and moving to the next period. Since the employer's problem is a standard one in dynamic programming, it is well-known that V^* attains a unique optimal value.

We may write the value function as

$$V^* = \delta \int_0^1 \max\left[H(q, V^*), V^*\right] dG(q) - k \tag{1}$$

where $H(q, V^*) = qv + (1 - q)(-c + V^*)$. Note that, according to our timing convention, cost k is incurred immediately while the payoff from hiring, $H(q, V^*)$, is received in the next period. Furthermore, since the employer's problem is stationary, any strategy attaining V^* must be a threshold strategy (see, for example, McCall, 1970). And the threshold must be the same for both kinds of candidates since, conditional on q, a candidate's kind γ is completely irrelevant: the only thing that matters is the probability of success itself, and not whether the candidate is mainstream or minority.

Under a generic threshold strategy, which we denote by $\underline{q}$, the value function given in Equation (1) reduces to

$$V\left(\underline{q}\right)=\delta\left[G\left(\underline{q}\right)V\left(\underline{q}\right)+\int_{\underline{q}}^{1}H\left(q,V\left(\underline{q}\right)\right)dG\left(q\right)\right]-k$$

Substituting for H and solving for $V\left(\underline{q}\right)$, we obtain

$$V\left(\underline{q}\right)=\frac{\delta\int_{\underline{q}}^{1}\left(qv+\left(1-q\right)\left(-c\right)\right)dG\left(q\right)-k}{1-\delta\left(1-\int_{\underline{q}}^{1}qdG\left(q\right)\right)}$$

Thus, the employer's problem reduces to choosing $\underline{q}$ to maximize $V\left(\underline{q}\right)$. Proposition 1 characterizes the unique optimum.

Proposition 1 *The optimal threshold strategy, $\underline{q}^*$, is the unique interior solution to*

$$\underline{q}^*=\frac{\left(1-\delta\left(1-\int_{\underline{q}^*}^{1}qdG\left(q\right)\right)\right)c}{\left(1-\delta G\left(\underline{q}^*\right)\right)c+\left(1-\delta\right)v+k} \tag{2}$$

Finally, Proposition 2 establishes that every possible threshold success probability can be an optimum.

Proposition 2 *For all $q\in[0,1)$, there exist parameter values such that $\underline{q}^*=q$.*

III. Performance Metrics

Recall that the optimal hiring strategy established in Proposition 1 is "color-blind" in the sense that the employer sets the same threshold success probability for both kinds of candidates. In this section, we study the implications of a uniform hiring threshold for observable performance metrics of diversity.

Permanent Workforce Composition

Perhaps the most important performance metric of diversity is the fraction of minorities in the permanent workforce of an organization, relative to their share in the underlying population. In terms of our model, this corresponds to the probability that a permanently hired candidate is a minority.

Formally, let r_γ denote the probability that the vacancy is permanently filled by a candidate of kind γ, when the employer uses the, not necessarily optimal, threshold strategy $\underline{q}$. Then, r_γ can be expressed recursively as follows.

$$\begin{aligned} r_\gamma &= m_\gamma \left(p\left(\left(1 - G_\gamma\left(\underline{q}|\Theta = 1\right)\right) + G_\gamma\left(\underline{q}|\Theta = 1\right) r_\gamma\right) + (1-p)\, r_\gamma\right) \\ &\quad + (1 - m_\gamma)\left(\left(1 - p\left(1 - G_{-\gamma}\left(\underline{q}|\Theta_{-\gamma} = 1\right)\right)\right) r_\gamma\right) \end{aligned}$$

We can write this expression much more compactly if we define $G_{\gamma\theta}$ to be the probability that a candidate of kind γ and type θ induces a posterior success probability less than or equal to $\underline{q}$. Formally, $G_{\gamma\theta} \equiv G_\gamma\left(\underline{q}|\Theta = \theta\right)$. Solving for r_γ, we obtain, in our more economical notation,

$$r_\gamma = \frac{m_\gamma\left(1 - G_{\gamma 1}\right)}{1 - m_\gamma G_{\gamma 1} - (1 - m_\gamma)\, G_{-\gamma 1}}$$

We want to compare minority representation in the workplace, r_B, with the minority share of the underlying population, m_B. Minorities are proportionally represented in the workplace when $\frac{r_\gamma}{m_\gamma} = 1$. It is easily verified that this is equivalent to the condition that $G_{A1} = G_{B1}$. In other words, minorities are proportionally represented if and only if the probability of type I error (rejection of competent candidates) is the same for both kinds of candidates. When does equality of Type I error hold?

Lemma 3 *There exists a unique threshold,* $\underline{q}^1 \equiv \frac{1}{1 + \frac{1-p}{p} e^{\frac{1}{2\sigma_A \sigma_B}}} < p$, *where the probability of type I errors is the same for both kinds of candidates.*

Unsurprisingly, the optimal threshold $\underline{q}^*$ given in Proposition 1 is generically not equal to $\underline{q}^1$. The next proposition shows that, depending on the relationship between $\underline{q}^*$ and $\underline{q}^1$, minorities may be under or overrepresented in the workplace.

Proposition 3

1. *Minorities are overrepresented in the workplace (i.e.,* $\frac{r_B}{m_B} > 1$*) iff* $0 < \underline{q}^* < \underline{q}^1$.
2. *Minorities are underrepresented (i.e.,* $\frac{r_B}{m_B} \leq 1$*) iff* $\underline{q}^1 \leq \underline{q}^* < 1$.

If $\underline{q}^1 < \underline{q}^*$, then minority candidates are more subject to Type I error than majority candidates. That is, competent minority candidates are rejected at a higher rate than competent majority candidates. This results in underrepresentation of minorities in the workplace relative to their share in the underlying population. On the other hand, if $\underline{q}^1 > \underline{q}^*$, then it is the majority candidates who are more subject to Type I error. This results in minority candidates being *over*represented in the workplace. Hence, the outcome depends on how "choosy" the employer is.

The following figure illustrates how the difference in Type I errors for majority and minority candidates varies with the threshold strategy of the employer. It displays the ratio $\frac{1-G_{A1}}{1-G_{B1}}$ of hiring probabilities for competent majority versus competent minority candidates as a function of the employer's threshold strategy $\underline{q}$. The parameter values used to draw the figure are: $p = .3$, $\sigma_A = 1$, $\sigma_B = \sqrt{2}$. Notice that at low thresholds $\left(\underline{q} < \underline{q}^1 < p\right)$ minority candidates are overrepresented in the workforce, and this disparity grows as the threshold increases from $\underline{q} = 0$. Since the workforce proportions exactly reflect those of the candidate population at $\underline{q} = \underline{q}^1$, minority overrepresentation must reverse itself for a sufficiently choosy employer. In the figure, the degree of minority overrepresentation is greatest at $\underline{q} = 0.18$ and declines thereafter. For thresholds $\underline{q} > \underline{q}^1$, the effect of the difference in type I errors can be quite severe for competent minority candidates. By the time the threshold reaches 0.7, a competent majority candidate stands an almost 140 times better chance of being hired than a competent minority candidate. Indeed, as the figure illustrates, the ratio of hiring probabilities increases without bound as the threshold approaches 1. [**Figure 1 Here**]

The figure illustrates that it becomes exceedingly unlikely that a minority candidate will fill the position as the threshold increases. Put differently, the workplace composition becomes increasingly homogeneous. As we show in the next proposition, the positive relationship between the choosiness of an employer and the homogeneity of the workplace is a general property of the model.

Proposition 4 *Suppose that the employer is "selective" in its hiring policy, i.e.,* $\underline{q} > p$*, then:*

1. As the employer becomes more selective, minority representation in the workplace decreases. Formally, r_B *is decreasing in* $\underline{q}$*.*

2. As the employer becomes arbitrarily selective, minorities vanish from the workplace. Formally, $\lim_{\underline{q}\to 1} r_B = 0$*.*

One may wonder what conditions on primitives guarantee that an employer will indeed be selective in the sense described in Proposition 4. A useful lower bound on the optimal threshold may be derived from the case of a "myopic" employer who only derives benefit one period into the future. Such an employer would choose a "break-even" threshold where $vq - (1-c)\,q = 0$ or equivalently, $q = \frac{c}{c+v}$. Employers who value payoffs in periods beyond the next will optimally raise the threshold above the break-even level to capture some of the option value of waiting. Hence, $\underline{q}^* > \frac{c}{c+v}$. As a result, a sufficient condition for an employer to be selective is that $p < \frac{c}{c+v}$.

Initial Hiring Rates

We have shown that differences in Type I errors can lead to gross differences between the share of minorities in the permanent workforce compared to their share of the candidate population. On the other hand, given the "color blind" threshold strategy of the employer, one might speculate that the fraction of minorities among initial hires *would* reflect the underlying population. As we shall see, this is not typically the case. Define the fraction of initial hires who are of kind γ as

$$h_\gamma = \frac{m_\gamma\,(1-G_\gamma)}{m_\gamma\,(1-G_\gamma) + m_{-\gamma}\,(1-G_{-\gamma})}$$

Notice that the probability that a candidate of kind γ will be hired, $1-G_\gamma$, consists of the probability of two separate events: (i) The joint event that the candidate is competent and passes the interview; (ii) The joint event that the candidate is incompetent and passes the interview. Event (ii) is equivalent to the probability of Type II error.

Having previously established a threshold, $\underline{q}^1$, where Type I error is equalized across the two kinds of candidates, it is useful to determine the analogous threshold where Type II error is equalized. That is, define $\underline{q}^0$ to be the threshold such that $G_{A0} = G_{B0}$, which has as its solution

$$\underline{q}^0 = \frac{1}{1 + \frac{1-p}{p} e^{-\frac{1}{2\sigma_A \sigma_B}}} > p$$

When $\underline{q} < \underline{q}^0$, notice that incompetent minority candidates have a greater chance of being hired than incompetent majority candidates, while for $\underline{q} > \underline{q}^0$ the opposite holds. Furthermore, notice that the threshold at which Type II error is equalized always lies above that where Type I error is equalized. That is, $\underline{q}^1 < \underline{q}^0$.

Finally, we turn our attention to the threshold, $\underline{q}^{\emptyset}$, where the initial hiring proportions are equal to the underlying population proportions. That is, $\underline{q}^{\emptyset}$ solves $G_A = G_B$. Unlike for the thresholds for equal Type I and Type II errors, there exists no closed-form solution for $\underline{q}^{\emptyset}$. However, from the fact that G_A is a mean-preserving spread of G_B (Lemma 2), it follows that $\underline{q}^{\emptyset}$ exists and is unique. Moreover, since $\underline{q}^{\emptyset}$ represents a trade-off between Type I and Type II errors, $\underline{q}^1 < \underline{q}^{\emptyset} < \underline{q}^0$.

As was the case for the composition of the permanent workforce, depending on the optimal threshold $\underline{q}^*$, minorities may be under or overrepresented among initial hires. Using arguments identical to those in Proposition 3, it may be readily shown that

Proposition 5

1. *Minorities are overrepresented among initial hires (i.e., $\frac{h_B}{m_B} > 1$) iff $0 < \underline{q}^* < \underline{q}^{\emptyset}$.*
2. *Minorities are underrepresented (i.e., $\frac{h_B}{m_B} \leq 1$) iff $\underline{q}^{\emptyset} \leq \underline{q}^* < 1$.*

It is interesting to note that, since $\underline{q}^1 < \underline{q}^{\emptyset}$, it may well be that an employer's optimal policy leads to favorable initial hiring rates for minorities, while their greater firing rates lead to underrepresentation in the permanent workforce. We turn to formally analyzing firing rates next.

Firing Rates

We saw that minority over- or underrepresentation among initial hires and in the permanent workforce depends on the threshold strategy of the employer. In the case of firing rates, by contrast, the model delivers unambiguous predictions: minority hires are fired at higher rates than majority hires for all (interior) threshold strategies $\underline{q} \in (0, 1)$.

The firing rate for hires of kind γ, which we denote by f_γ, is equal to the probability that a candidate of kind γ is incompetent conditional on actually having been hired. Formally,

$$\begin{aligned} f_\gamma &= \Pr\left(\Theta = 0 | Q_\gamma \geq \underline{q}\right) \\ &= \frac{\left(1 - G_{\gamma 0}\right)\left(1 - p\right)}{1 - G_\gamma} \end{aligned}$$

A simple intuition might suggest that firing rates simply reflect Type II errors in the screening decision, i.e. $\Pr\left[Q_\gamma \geq \underline{q} | \Theta = 0\right]$. Indeed, the firing rate is very much affected by Type II error. However, the two are by no means the same. The reason is that the pool of initial hires from which people are fired also depends on the ex ante probability of *false negatives* in the subpopulation—i.e., Type I error. Thus, both types of error interact to produce the firing rate of a subpopulation. To see how firing rates reflect the trade-off between Type I and Type II errors, it is helpful to write f_γ as follows

$$f_\gamma = \frac{(1-p)\Pr(\text{Type II})}{(1-p)\Pr(\text{Type II}) + p\left(1 - \Pr(\text{Type I})\right)}$$

When $\underline{q}^1 \leq \underline{q} \leq \underline{q}^0$, minorities suffer greater Type I and Type II errors than do majorities. As a consequence, the firing rate of minorities is higher than for majorities. When $\underline{q} < \underline{q}^1$, minorities continue to experience greater Type II error; however, Type I error is now higher for majorities than for minorities. As a consequence, the ordering of majority and minority firing rates becomes ambiguous and depends on the relative magnitude of the two types of errors. Similarly, when $\underline{q} > \underline{q}^0$, Type II error is smaller for minorities than for majorities but Type I error is greater. Hence, also in this case, the ordering could go either way. As the next proposition shows, however, the trade-off between Type I and Type II errors is always resolved in the direction of higher firing rates for minorities.[3]

Proposition 6 *For all $\underline{q} \in (0,1)$, minority hires are fired at a higher rate than majority hires.*

Summary

The following figure summarizes the various performance metrics of diversity as a function of the success probability threshold $\underline{q}$. **[Figure 2 Here]**

[3]Proposition 6 ignores the cases where $\underline{q} \in \{0,1\}$ since, for these degenerate cases, either everyone is hired or no one is hired and the firing rate problem is therefore trivial.

IV. Policy Implications

In this section, we examine how the optimal threshold—and, by implication, the diversity metrics described above—vary with changes in the parameters of the model. Some of these parameters are likely to be under policy control; hence, there is the possibility of influencing workplace diversity. Throughout this section, we shall use the term workplace diversity as being synonymous with the minority representation ratio $\frac{r_B}{m_B}$. The closer this ratio is to one, the more diverse is the workplace. Also, we shall assume that the employer is "selective" in its hiring policy, i.e., $\underline{q}^* > p$. Therefore, an increase in $\underline{q}^*$ implies a decrease in diversity, and vice versa.

Diversity and Worker Protections

There has been considerable debate, especially in Europe, over the appropriate level of worker protections against dismissal. The mass street protests in France during the Spring of 2006 against the "contrat première embauche" are a salient example. This new law would have allowed for summary dismissal of employees below the age of 26 during the first two years of their contract. By reducing the risk of hiring, it was hoped that the contrat première embauche would lead to a reduction in the very high youth unemployment. Whether it would have achieved its goal shall remain unknown, as the law was retracted in response to the protests. In terms of our model, European-style worker protection policies may be thought of as increasing the cost of firing, c.

Implication 1 *An increase in the cost of firing, c, reduces workplace diversity.*

Intuitively, raising the cost of firing increases the cost of Type II errors for the employer. As a result, he becomes more reluctant to take a chance on whether a candidate can do the job and, consequently, raises the threshold for hiring. As we have shown in the previous section, when the employer is at all selective, increased hiring thresholds have the effect of differentially raising Type I errors to the disadvantage of minorities. As a result, workplace diversity decreases.

Diversity and the Cost of Recruiting

Video conferencing over the internet is decreasing firms' costs of interviewing. (See, e.g., Matt Bolch, 2007.) This allows recruiters to conduct more interviews and, thereby, "widen their net." A common intuition suggests that such a widening would increase diversity in the workplace. The model, however, shows how this intuition can go wrong.

Implication 2 *A decrease in the cost of interviewing, k, reduces workplace diversity.*

By reducing the cost of interviewing, it becomes less expensive for the employer to be choosy. As a result, the employer raises his threshold for hiring. This, in turn, increases the difference in Type I errors between minorities and majorities and, hence, reduces diversity. Also, notice that a reduction in k and a reduction in c both lower frictions in hiring. Nevertheless, they have opposite effects on diversity: while an reduction in c raises workplace diversity, a reduction in k lowers it.

Diversity over the Business Cycle

Next, we consider how the employer's optimal threshold varies with the business cycle.[4] At a peak in the business cycle, job candidates become more scarce and, hence, the cost of recruiting increases. As we have shown above, this has the effect of raising workplace diversity. In addition, the value-added of a competent employee is also likely to be higher at the peak of the business cycle than during a recession. In terms of our model, this corresponds to an increase in v.

Implication 3 *Diversity is procyclical. Formally, $\underline{q}^*$ is decreasing in v (and k).*

Intuitively, as a competent employee's value-added increases, it becomes more costly to leave the position unfilled. As a consequence, the employer is more willing to take a chance by hiring possibly incompetent employees and, hence, the optimal threshold falls. The lower threshold reduces the difference in Type I errors between minorities and majorities. Consequently, workplace diversity increases. As mentioned in the introduction, the procyclicality of diversity is indeed consistent with the Dutch experience over the last decade.

[4] The analysis here uses only the comparative static implications of the stationary model to derive conclusions. This is appropriate if firms view changes in the business climate to be permanent rather than transitory. Failing this, one needs to account for the non-stationarity of the future value of not hiring conditional on the present state of the economy—a considerably more involved dynamic programming problem that is beyond the scope of this paper.

Diversity and the Cost of Capital

[[**John will actually REMEMBER to ask Hayne Leland or other finance guys about this.**]] Another testable implication of the model is that variation in the riskiness of firms leads to differences in workplace diversity. If we interpret the discount parameter δ as representing an employer's cost of capital, which presumably varies with the riskiness of his business, then we have the following implication:

Implication 4 *Riskier firms are more diverse. Formally, $\underline{q}^*$ is increasing in δ.*

Intuitively, the option value of waiting is worth less for risky firms than for safe firms. Since the degree to which the optimal threshold lies above the break-even threshold positively depends on this option value, the optimal threshold for a riskier firm is lower than that for a less risky firm. In turn, this lower threshold reduces the difference in Type I errors between minorities and majorities, and, consequently, workplace diversity increases.

Diversity and the Scarcity of Competence

As we highlighted above, the key determinant of minority over- or underrepresentation is the relationship between the optimal threshold $\underline{q}^*$ and the thresholds for equating Type I and Type II errors across the two populations—$\underline{q}^1$ and $\underline{q}^0$, respectively. These two thresholds bracket the prior probability that a candidate is competent; that is, $\underline{q}^1 < p < \underline{q}^0$. At the same time, the thresholds are a function of p. Thus, a question that naturally arises is how under- or overrepresentation varies with the underlying probability that a candidate can do the job.

When few candidates can do the job, i.e., when p is low, the results of the interview must be sufficiently convincing to induce the employer to take a chance on the candidate given the costs of firing. A candidate with a very noisy signal is going to have a difficult time in making this case. In the limit, imagine a situation where B candidates have arbitrarily noisy signals and where the employer is selective. Clearly, there is virtually no possibility of overcoming the employer's prior belief about the low likelihood that the candidate is qualified. In contrast, a candidate with a very precise signal faces no such handicap. In this extreme case, one would expect (and the model predicts) severe underrepresentation of minority candidates both at the hiring stage and in the permanent workforce.

By contrast, when most candidates can do the job, i.e., when p is high, an imprecise signal in the interview stage can be an advantage for a candidate. Suppose that p is so high such that the employer is predisposed to give most candidates a chance to prove themselves on the job. In that case, having an arbitrarily noisy signal virtually guarantees that the candidate will not greatly disappoint in the interview and, hence, that he will be offered the position. In contrast, a more precise signal exposes the candidate to a greater possibility of making a bad impression in the interview and, hence, of being declined the job—even in the case where the candidate is in fact competent. In this situation, overrepresentation of minority candidates, both in hiring and in the permanent workforce, is the more likely outcome. The next implication formalizes this intuition.

Implication 5 *In jobs that require exceptional skills, minorities will be underrepresented. In jobs that require very common skills, minorities will be overrepresented. Formally, there exists* $0 < p_0 < p_1 < 1$ *such that, for all* $p \in (0, p_0)$, $\frac{r_B}{m_B} < 1$, *while for all* $p \in (p_1, 1)$, $\frac{r_B}{m_B} > 1$.

Diversity and the Size of the Minority Population

In recent years Europe's minority population has grown rapidly, owing both to immigration and higher fertility rates. Will this change in the relative composition of the employee pool improve diversity in the workplace? In terms of the model, this question amounts to a study of the comparative static implications of an increase in m_B, the minority fraction of the population.

Implication 6 *The larger the minority fraction of the population, the smaller its degree of underrepresentation in the workforce. Formally,* $\underline{q}^*$ *is decreasing in* m_B.

Intuitively, because of the non-directed nature of the search process, an increase in the minority fraction of the population leads the employer to (optimally) become less selective in its hiring decisions (since the expected time to a hire under a given level of selectivity has now gone up). As a consequence, the fraction of minorities among initial hires and in the permanent workforce increases. A similar result can be found in Lundberg and Startz (forthcoming).

V. Conclusion

In this paper we have investigated the implications of the assumption that employers find it easier to accurately evaluate majority job candidates than minority job candidates. We have shown that this basic premise implies that there exists a tension between job security and workplace diversity. When job security is high, that is, firing non-performing staff is expensive, minorities are likely to be severely underrepresented in the workplace, particularly in demanding positions. At the other extreme the converse holds; when job security is low, minorities are overrepresented in undemanding positions. These distortions occur even though majority and minority populations have identical skill levels.

On a fundamental level, our results are driven by Bayes' law, which implies that employers' posterior beliefs about majority candidates respond more strongly to new information than their beliefs about minority candidates. When the information received is better than expected, this high belief-sensitivity works to the advantage of majority candidates. On the other hand, when the information is worse than expected, high belief-sensitivity works to the *dis*advantage of majority candidates.

While the occurrence of "reverse discrimination" may be interesting from a theoretical perspective, from a policy perspective, the under representation of minorities in demanding positions seems the more relevant model prediction. Given that minorities are indeed grossly underrepresented at the higher levels of many organizations, what can be done about it?

In our model, the lack of workplace diversity arises owing to a communication mismatch between the majority employer/interviewer and minority job candidates. Obviously, matching the background of the interviewer with the background of the candidate would solve this problem. However, more often than not, this may not be feasible. A more realistic option to increase workplace diversity is to lower firing costs which, in turn, induces employers to be less choosy in the initial screening and creates an opportunity for competent minority candidates to prove themselves on the job.

There are several limitations in our modeling approach worth discussing. From a technical standpoint, one limitation is the one-sided search, or partial equilibrium nature of the analysis. It would be useful to extend the model to a general equilibrium framework. Also, the binary nature of competence—candidates either can do the job or they cannot—is clearly restrictive. Other limitations are of a less technical nature, such as the assumptions of equal average skill levels, identical firing costs across subpopulations, and no naked racism or directed search on the part of employers. Also, we have assumed that employers only care about technical competence, and not how a candidate fits into the culture of the organization. Some or all of these assumptions do not hold in practice. However, most realistic deviations all point in the same direction: towards more rather than less discrimination. As such, the model puts a lower bound on the problem and shows that, even under the best of circumstances, competent minority candidates are likely to have a much harder time securing a coveted job than equally competent majority candidates, in particular when job security is high.

Appendices

A. Proofs of Lemmas

Lemma 1 *For all $p > p'$, $G(\cdot\,; p)$ first-order stochastically dominates $G(\cdot\,; p')$. That is,*

$$\frac{d}{dp}G(q) < 0, \text{ for all } q \in (0,1)$$

Proof. Recall that

$$G(q) = (1 - m_B)\, G_A(q) + m_B G_B(q)$$

where

$$G_\gamma(q) = p\Phi\left(\frac{s_\gamma(q) - 1}{\sigma_\gamma}\right) + (1-p)\,\Phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right)$$

$\gamma = A, B$.

Now, $\frac{d}{dp}G_\gamma(q) =$

$$= \Phi\left(\frac{s_\gamma(q) - 1}{\sigma_\gamma}\right) - \Phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right) + \left(p\phi\left(\frac{s_\gamma(q) - 1}{\sigma_\gamma}\right) + (1-p)\,\phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right)\right)\frac{\partial s_\gamma(q)}{\partial p} < 0$$

because $\frac{\partial s_\gamma(q_\gamma)}{\partial p} = -\frac{\sigma_\gamma^2}{p(1-p)} < 0$ and $\Phi\left(\frac{s_\gamma(q)-1}{\sigma_\gamma}\right) < \Phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right)$.

Since $G(q)$ is a convex combination of $G_A(q)$ and $G_B(q)$, it follows that $\frac{d}{dp}G(q) < 0$ for all $q \in (0,1)$. This proves the lemma. ■

Lemma 2 *$G_A(\cdot)$ is a mean preserving spread of $G_B(\cdot)$. And, for all $m_B < m_B'$, $G(\cdot\,; m_B)$ is a mean preserving spread of $G(\cdot\,; m_B')$.*

Proof. First, we verify that $E_{G_A}[Q_A] = E_{G_B}[Q_B] = p$.

By definition,

$$E_{G_\gamma}[Q_\gamma] = \int_0^1 q g_\gamma(q)\, dq$$

where $\gamma \in \{A, B\}$. Changing the integration variable from probability q to signal s, we get

$$E_{G_\gamma}[Q_\gamma] = \int_{-\infty}^{\infty} q_\gamma(s)\, g_\gamma(s)\, \frac{dq_\gamma(s)}{ds} ds$$

where $q_\gamma(s) = \frac{p\phi\left(\frac{s-1}{\sigma_\gamma}\right)}{p\phi\left(\frac{s-1}{\sigma_\gamma}\right)+(1-p)\phi\left(\frac{s}{\sigma_\gamma}\right)}$, $\frac{\partial q_\gamma(s)}{\partial s} = \frac{p(1-p)}{\sigma_\gamma}\frac{\phi\left(\frac{s}{\sigma_\gamma}\right)\phi\left(\frac{s-1}{\sigma_\gamma}\right)}{\left(p\phi\left(\frac{s-1}{\sigma_\gamma}\right)+(1-p)\phi\left(\frac{s}{\sigma_\gamma}\right)\right)^2}$ and $g_\gamma(s) = \left(p\phi\left(\frac{s-1}{\sigma_\gamma}\right) + (1-p)\phi\left(\frac{s}{\sigma_\gamma}\right)\right)\frac{\sigma_\gamma}{q_\gamma(s)(1-q_\gamma(s))}$. Hence,

$$\begin{aligned} E_{G_\gamma}[Q_\gamma] &= \int_{-\infty}^{\infty} q_\gamma(s)\, g_\gamma(s)\frac{dq_\gamma(s)}{ds}ds \\ &= p\int_{-\infty}^{\infty}\phi\left(\frac{s-1}{\sigma_\gamma}\right)ds \\ &= p \end{aligned}$$

This proves that $E_{G_A}[Q_A] = E_{G_B}[Q_B] = p$. For later use, note that $E_{G(\cdot;m_B)}[Q] = E_{G\left(\cdot;m'_B\right)}[Q] = p$.

To prove that $G_A(\cdot)$ is a mean preserving spread of $G_B(\cdot)$ it now suffices to show that, on the interval $(0,1)$, $G_B(\cdot)$ crosses $G_A(\cdot)$ only once and from below. We do this by establishing that the difference $D(q) \equiv G_A(q) - G_B(q)$ has two extrema: starting from zero at $q = 0$, $D(q)$ first reaching a maximum—at which $D(q)$ is strictly positive—and then a minimum—at which $D(q)$ is strictly negative.

Let

$$\zeta = \ln\left(\frac{1-q}{q}\frac{p}{1-p}\right)$$

such that

$$\begin{aligned} D &= G_A(q) - G_B(q) \\ &= p\Phi\left(\frac{-\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right) + (1-p)\Phi\left(\frac{\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right) \\ &\quad -p\Phi\left(\frac{-\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right) - (1-p)\Phi\left(\frac{\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right) \end{aligned}$$

Relying on the fact that ζ is a monotone function of q, we now ask when $\frac{dD}{d\zeta} = 0$:

$$\begin{aligned} \frac{dD}{d\zeta} &= -\sigma_A p\phi\left(\frac{-\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right) - \sigma_A(1-p)\phi\left(\frac{\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right) \\ &\quad +\sigma_B p\phi\left(\frac{-\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right) + \sigma_B(1-p)\phi\left(\frac{\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right) \\ &= 0 \end{aligned}$$

$\Longleftrightarrow$

$$\begin{aligned}
\frac{\sigma_A}{\sigma_B} &= \frac{\phi\left(\frac{-\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right)+\frac{1-p}{p}\phi\left(\frac{\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right)}{\phi\left(\frac{-\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right)+\frac{1-p}{p}\phi\left(\frac{\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right)} \\
&= \frac{e^{-\frac{1}{2}\left(\frac{-\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right)^2}+\frac{1-p}{p}e^{-\frac{1}{2}\left(\frac{\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right)^2}}{e^{-\frac{1}{2}\left(\frac{-\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right)^2}+\frac{1-p}{p}e^{-\frac{1}{2}\left(\frac{\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right)^2}}
\end{aligned}$$

Now consider the right-hand side, which we denote by Ψ, as a function of ζ.

$$\Psi \equiv \frac{e^{-\frac{1}{2}\left(\frac{-\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right)^2}+\frac{1-p}{p}e^{-\frac{1}{2}\left(\frac{\frac{1}{2}-\sigma_B^2\zeta}{\sigma_B}\right)^2}}{e^{-\frac{1}{2}\left(\frac{-\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right)^2}+\frac{1-p}{p}e^{-\frac{1}{2}\left(\frac{\frac{1}{2}-\sigma_A^2\zeta}{\sigma_A}\right)^2}}$$

$$\begin{aligned}
&= e^{-\frac{1}{2}\left(\frac{1}{4\sigma_B^2}+\sigma_B^2\zeta^2\right)+\frac{1}{2}\left(\frac{1}{4\sigma_A^2}+\sigma_A^2\zeta^2\right)} \\
&= e^{\left(\frac{1}{8}\left(\frac{1}{\sigma_A^2}-\frac{1}{\sigma_B^2}\right)+\frac{1}{2}\left(\sigma_A^2-\sigma_B^2\right)\zeta^2\right)}
\end{aligned}$$

Thus, D takes on extrema at values of ζ that solve

$$\frac{\sigma_A}{\sigma_B} = e^{\left(\frac{1}{8}\left(\frac{1}{\sigma_A^2}-\frac{1}{\sigma_B^2}\right)+\frac{1}{2}\left(\sigma_A^2-\sigma_B^2\right)\zeta^2\right)}$$

Taking logs

$$\ln\frac{\sigma_A}{\sigma_B} = \frac{1}{8}\left(\frac{1}{\sigma_A^2}-\frac{1}{\sigma_B^2}\right)+\frac{1}{2}\left(\sigma_A^2-\sigma_B^2\right)\zeta^2$$

Therefore, the solutions to ζ are roots of the function

$$\frac{1}{8}\left(\frac{1}{\sigma_A^2}-\frac{1}{\sigma_B^2}\right)+\frac{1}{2}\left(\sigma_A^2-\sigma_B^2\right)\zeta^2-\ln\frac{\sigma_A}{\sigma_B}$$

These roots are

$$\zeta = \frac{-1}{2\sigma_A\sigma_B\ln\frac{\sigma_B}{\sigma_A}},\ \zeta = \frac{1}{2\sigma_A\sigma_B\ln\frac{\sigma_B}{\sigma_A}}$$

The existence of exactly two distinct roots for ζ (and hence for q) implies that G_A and G_B cross each other exactly once. It remains to verify that G_B crosses G_A from below and not from above. Now,

$$\begin{aligned}
D &= G_A(q)-G_B(q) \\
&= p\left(G_{A1}-G_{B1}\right)+(1-p)\left(G_{A0}-G_{B0}\right)
\end{aligned}$$

At $q = \underline{q}^1 = \frac{1}{1+\frac{1-p}{p}e^{\frac{1}{2\sigma_A\sigma_B}}}$, $G_{A1} - G_{B1} = 0$ while $G_{A0} - G_{B0} > 0$. Hence, $D\left(\underline{q}^1\right) > 0$.

At $q = \underline{q}^0 = \frac{1}{1+\frac{1-p}{p}e^{-\frac{1}{2\sigma_A\sigma_B}}}$, $G_{A0} - G_{B0} = 0$ while $G_{A1} - G_{B1} < 0$. Hence, $D\left(\underline{q}^0\right) < 0$.

Now, because $\underline{q}^1 < \underline{q}^0$, this implies that G_B crosses G_A from below.

This completes the proof that $G_A(\cdot)$ is a mean-preserving spread of $G_B(\cdot)$.

Finally, to prove that $G(\cdot; m_B)$ is a mean preserving spread of $G(\cdot; m'_B)$ for all $m_B < m'_B$, it remains to show that $G(\cdot; m_B)$ second-order stochastically dominates $G(\cdot; m'_B)$. Or,

$$\int_0^{\hat{q}} G(q, m_B)\, dq - \int_0^{\hat{q}} G(q, m'_B)\, dq \leq 0$$

for all $\hat{q} \in (0,1)$, with strict inequality for some $\hat{q}$. Now,

$$\begin{aligned} & \int_0^{\hat{q}} G(q, m_B)\, dq - \int_0^{\hat{q}} G(q, m'_B)\, dq \\ = \quad & (m'_B - m_B) \int_0^{\hat{q}} (G_A(q) - G_B(q))\, dq \leq 0 \end{aligned}$$

where the weak inequality for all $\hat{q}$, and the strict inequality for some $\hat{q}$, follow from the fact that $G_B(\cdot)$ second-order stochastically dominates $G_A(\cdot)$.

This completes the proof. ■

Lemma 3 *There exists a unique threshold,* $\underline{q}^1 \equiv \frac{1}{1+\frac{1-p}{p}e^{\frac{1}{2\sigma_A\sigma_B}}} < p$, *where the probability of type I errors is the same for both kinds of candidates.*

Proof.

$$G_{A1}\left(\underline{q}\right) = G_{B1}\left(\underline{q}\right)$$

$\Leftrightarrow$

$$\Phi\left(\frac{s_A\left(\underline{q}\right) - 1}{\sigma_A}\right) = \Phi\left(\frac{s_B\left(\underline{q}\right) - 1}{\sigma_B}\right)$$

$\Leftrightarrow$

$$\frac{s_A\left(\underline{q}\right) - 1}{\sigma_A} = \frac{s_B\left(\underline{q}\right) - 1}{\sigma_B}$$

where

$$s_\gamma(q) = \frac{1}{2} - \sigma_\gamma^2 \ln\left(\frac{1-q}{q}\frac{p}{1-p}\right)$$

Hence,

$$\frac{\frac{1}{2} - \sigma_A^2 \ln\left(\left(\frac{1}{\underline{q}} - 1\right)\frac{p}{1-p}\right) - 1}{\sigma_A} = \frac{\frac{1}{2} - \sigma_B^2 \ln\left(\left(\frac{1}{\underline{q}} - 1\right)\frac{p}{1-p}\right) - 1}{\sigma_B}$$

$$\sigma_A Z + \frac{1}{2\sigma_A} = \sigma_B Z + \frac{1}{2\sigma_B}$$

$$\begin{aligned}
\sigma_B \sigma_A^2 \ln\left(\left(\frac{1}{\underline{q}} - 1\right)\frac{p}{1-p}\right) + \frac{1}{2}\sigma_B &= \sigma_A \sigma_B^2 \ln\left(\left(\frac{1}{\underline{q}} - 1\right)\frac{p}{1-p}\right) + \frac{1}{2}\sigma_A \\
Z &= \frac{1}{2\sigma_B \sigma_A}
\end{aligned}$$

$\Leftrightarrow$

$$\underline{q} = \frac{1}{1 + \frac{1-p}{p} e^{\frac{1}{2\sigma_A \sigma_B}}}$$

■

Lemma 4 *Suppose $\underline{q} > p$. Then:*

1. *The distribution G_{A1} dominates G_{B1} in terms of the likelihood ratio.*
2. *The distribution G_{A0} dominates G_{B0} in terms of the likelihood ratio.*

Proof. To establish this, it is sufficient to show that $\frac{\partial^2 \ln g_{A1}}{\partial\sigma\partial q} > 0$.

$$\begin{aligned}
\frac{\partial^2 \ln g_{A1}}{\partial\sigma\partial q} &= \frac{\partial^2 \ln \phi\left(\frac{s(q)-1}{\sigma}\right)\frac{\sigma}{q(1-q)}}{\partial\sigma\partial q} \\
&= \frac{\partial^2 \ln\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{s(q)-1}{\sigma}\right)^2} \frac{\sigma}{q(1-q)}\right)}{\partial\sigma\partial q} \\
&= \frac{2}{q}\sigma \frac{\ln\left(\frac{1-q}{q}\frac{p}{1-p}\right)}{1-q} > 0
\end{aligned}$$

where the inequality holds since $q > p$. The proof of part 2 of the Lemma is virtually identical. ■

Lemma 5 *Suppose $\underline{q} > p$. Then:*

1. *The distribution G_{A1} dominates G_{B1} in terms of the hazard rate.*
2. *The distribution G_{A0} dominates G_{B0} in terms of the hazard rate.*

Proof. Lemma 4 implies that

$$\frac{g_{B1}(q')}{g_{B1}(q)} < \frac{g_{A1}(q')}{g_{A1}(q)}$$

for all $p < q < q'$.

Hence,

$$\int_q^1 \frac{g_{A1}(t)}{g_{A1}(q)}dt > \int_q^1 \frac{g_{B1}(t)}{g_{B1}(q)}$$
$$\frac{1-G_{A1}(q)}{g_{A1}(q)} > \frac{1-G_{B1}(q)}{g_{B1}(q)}$$

or, equivalently,

$$\frac{g_{A1}(q)}{1-G_{A1}(q)} < \frac{g_{B1}(q)}{1-G_{B1}(q)}$$

The proof of part 2 of the lemma is virtually identical. ■

B. Proofs of Propositions

Proposition 1 *The optimal threshold, $\underline{q}^*$, is the unique interior solution to*

$$\underline{q}^* = \frac{\left(1-\delta\left(1-\int_{\underline{q}^*}^1 qdG(q)\right)\right)c}{\left(1-\delta G\left(\underline{q}^*\right)\right)c+(1-\delta)v+k}$$

Proof. Recall that

$$\begin{aligned} V\left(\underline{q}\right) &= \frac{\delta\int_{\underline{q}}^1 (qv+(1-q)(-c))\,dG(q)-k}{1-\delta\left(1-\int_{\underline{q}}^1 qdG(q)\right)} \\ &= \frac{\delta v\int_{\underline{q}}^1 qdG(q)-\delta c\left(1-G\left(\underline{q}\right)\right)+\delta c\int_{\underline{q}}^1 qdG(q)-k}{1-\delta\left(1-\int_{\underline{q}}^1 qdG(q)\right)} \end{aligned}$$

It is useful to represent this as numerator and denominator components for purposes of differentiation. Hence, define

$$N \equiv \delta\int_{\underline{q}}^1 (qv+(1-q)(-c))\,dG(q)-k$$

and

$$D \equiv 1-\delta\left(1-\int_{\underline{q}}^1 qdG(q)\right)$$

Thus, the first-order necessary condition for optimality, $\frac{\partial V(\underline{q})}{\partial \underline{q}}=0$, may be expressed as

$$\frac{DN'-ND'}{D^2}=0$$

Therefore,

$$\begin{aligned}
\frac{\partial V(\underline{q})}{\partial \underline{q}} &= \frac{D\left(-\delta g(\underline{q})\left((v+c)\underline{q}-c\right)\right)-N\left(-\delta \underline{q} g(\underline{q})\right)}{D^2} \\
&= \delta g(\underline{q}) \frac{-D(v+c)\underline{q}+Dc+N\underline{q}}{D^2}
\end{aligned}$$

Hence,

$$-D(v+c)\underline{q}+Dc+N\underline{q}=0$$

and this implies that

$$\underline{q}^* = \frac{Dc}{D(v+c)-N}$$

Substituting for D and N, and simplifying, we get the following implicit characterization of $\underline{q}^*$:

$$\begin{aligned}
\underline{q}^* &= \frac{\left(1-\delta\left(1-\int_{\underline{q}^*}^1 q dG(q)\right)\right)c}{\left(1-\delta\left(1-\int_{\underline{q}^*}^1 q dG(q)\right)\right)(v+c)-\delta\int_{\underline{q}^*}^1 (qv+(1-q)(-c))\,dG(q)+k} \\
&= \frac{\left(1-\delta\left(1-\int_{\underline{q}^*}^1 q dG(q)\right)\right)c}{\left(1-\delta G\left(\underline{q}^*\right)\right)c+(1-\delta)v+k}
\end{aligned}$$

and this yields the expression in Lemma 1.

Having derived the necessary first-order condition for an interior solution $\underline{q}^* \in (0,1)$, we now prove its actual existence.

At $\underline{q}^*=0$, LHS $<$ RHS. At $\underline{q}^*=1$, LHS $>$ RHS. Hence, by continuity and the intermediate value theorem, there must be a $\underline{q}^* \in (0,1)$ such that LHS $=$ RHS.

Next, we prove uniqueness by showing that there is at most one $\underline{q}^* \in (0,1)$ that satisfies the necessary first-order condition.

To see this, first notice that we may rewrite the first-order condition as follows:

$$\underline{q}^*(c+(1-\delta)v+k) = c - c\delta\left(1-\int_{\underline{q}^*}^1 q dG(q)\right)+\delta G\left(\underline{q}^*\right)c\underline{q}^*$$

Integrating by parts, we obtain

$$\underline{q}^*(c+(1-\delta)v+k) = c - c\delta\int_{\underline{q}^*}^1 G(q)\,dq$$

Adding and subtracting $c\delta \int_0^{\underline{q}^*} G(q)\,dq$ to the right-hand side yields

$$\underline{q}^*(c + (1-\delta)v + k) = c - c\delta \int_0^1 G(q)\,dq + c\delta \int_0^{\underline{q}^*} G(q)\,dq$$

Finally, noting that $\int_0^1 G(q)\,dq = 1 - p$ and substituting, we obtain

$$\underline{q}^*(c + (1-\delta)v + k) = c(1-\delta) + c\delta \left(p + \int_0^{\underline{q}^*} G(q)\,dq \right)$$

Hence,

$$\underline{q}^* = \frac{(1-\delta)c + c\delta p}{(c + (1-\delta)v + k)} + \frac{c\delta}{(c + (1-\delta)v + k)} \int_0^{\underline{q}^*} G(q)\,dq$$

Note that the right-hand side is monotonically increasing in $\underline{q}^*$ at a speed < 1, for all $\underline{q}^* \in (0,1)$. This implies, however, that the right-hand side can cross the 45-degree line, which corresponds to the left-hand side, at most once. Hence, there is at most one $\underline{q}^* \in (0,1)$ that satisfies the necessary first-order condition.

Finally, we show that at the unique interior $\underline{q}^*$, the value function reaches a global maximum. This follows from the observation that $\lim_{\underline{q}\to 1} V(\underline{q}) \to -\infty$, and that there exists an $\varepsilon > 0$ such that for all $0 < \underline{q} < \varepsilon$, $\frac{\partial V(\underline{q})}{\partial \underline{q}} > 0$. To see that the latter assertion is indeed true, recall that

$$V(\underline{q}) = \frac{\delta \int_{\underline{q}}^1 (qv + (1-q)(-c))\,dG(q) - k}{1 - \delta \left(1 - \int_{\underline{q}}^1 q\,dG(q)\right)}$$

and that

$$\frac{\partial V(\underline{q})}{\partial \underline{q}} = \delta g(\underline{q}) \frac{-D(v+c)\underline{q} + Dc + N\underline{q}}{D^2}$$

where N and D denote the numerator and the denominator of $V(\underline{q})$, respectively.

Now we rewrite $\frac{\partial V(\underline{q})}{\partial \underline{q}}$ to get

$$\frac{\partial V(\underline{q})}{\partial \underline{q}} = \delta g(\underline{q}) \left(\frac{c}{D} + \frac{V(\underline{q}) - (v+c)}{D} \underline{q} \right)$$

Written in this form, it is obvious that, for sufficiently small $\underline{q} > 0$, both factors in the last expression are strictly positive. This proves the proposition. ■

Proposition 2 *For all $q \in [0,1)$, there exist parameter values such that $\underline{q}^* = q$.*

Proof. Fix $k = 0$. In that case, the employer will always wish to participate by interviewing candidates rather than eschewing the employment market. When $c = 0$, the right-hand side of equation (2) equals zero; hence, $\underline{q}^* = 0$. When $c \to \infty$, the right-hand side of equation (2) goes to 1 as the following argument shows:

$$\begin{aligned} & \lim_{c\to\infty} \frac{\left(1-\delta\left(1-\int_{\underline{q}^*}^1 q dG(q)\right)\right)c}{\left(1-\delta G\left(\underline{q}^*\right)\right)c+(1-\delta)v} \\ \geq \;& \lim_{c\to\infty} \frac{\left(1-\delta\left(1-\int_{\underline{q}^*}^1 dG(q)\right)\right)c}{\left(1-\delta G\left(\underline{q}^*\right)\right)c+(1-\delta)v} \\ = \;& \lim_{c\to\infty} \frac{\left(1-\delta G\left(\underline{q}^*\right)\right)c}{\left(1-\delta G\left(\underline{q}^*\right)\right)c+(1-\delta)v} = 1 \end{aligned}$$

Hence, $\lim_{c\to\infty} \underline{q}^* = 1$. Finally, since the right-hand side of equation (2) is continuous in c, it follows that there exist parameter values such that $\underline{q}^* = q$ for all $q \in [0, 1)$. ■

Proposition 3

1. *Minorities are overrepresented in the workplace (i.e.,* $\frac{r_B}{m_B} > 1$*) iff* $0 < \underline{q}^* < \underline{q}^1$*.*
2. *Minorities are underrepresented (i.e.,* $\frac{r_B}{m_B} \leq 1$*) iff* $\underline{q}^1 < \underline{q}^* < 1$*.*

Proof. Under a uniform threshold success probability $\underline{q}$, $\frac{r_B}{m_B} = 1$ iff $G_{A1}\left(\underline{q}\right) = G_{B1}\left(\underline{q}\right)$. As we saw in Lemma 3, this corresponds to $\underline{q} = \underline{q}^1 = \frac{1}{1+\frac{1-p}{p}e^{\frac{1}{2\sigma_A\sigma_B}}}$. To prove the proposition, we show that at the critical point $\underline{q}^1$, raising $\underline{q}$ leads to strict underrepresentation of minorities. That is, we calculate the derivative of

$$G_{A1}(q) - G_{B1}(q) = \Phi\left(\frac{s_A(q)-1}{\sigma_A}\right) - \Phi\left(\frac{s_B(q)-1}{\sigma_B}\right)$$

with respect to q, evaluate it at $\underline{q}^1 = \frac{1}{1+\frac{1-p}{p}e^{\frac{1}{2\sigma_A\sigma_B}}}$ and show that it is strictly negative.

The derivative is equal to

$$g_{A1}(q) - g_{B1}(q) = \phi\left(\frac{s_A(q)-1}{\sigma_A}\right)\frac{\sigma_A}{q(1-q)} - \phi\left(\frac{s_B(q)-1}{\sigma_B}\right)\frac{\sigma_B}{q(1-q)}$$

Multiplying by $q(1-q)$ and evaluating at $\underline{q}^1$, we get

$$\begin{aligned} =\;& \phi\left(\frac{\frac{1}{2}\frac{\sigma_B-\sigma_A}{\sigma_B}-1}{\sigma_A}\right)\sigma_A - \phi\left(\frac{-\frac{1}{2}\frac{\sigma_B-\sigma_A}{\sigma_A}-1}{\sigma_B}\right)\sigma_B \\ =\;& \phi\left(-\frac{1}{2}\frac{\sigma_B+\sigma_A}{\sigma_A\sigma_B}\right)\sigma_A - \phi\left(-\frac{1}{2}\frac{\sigma_B+\sigma_A}{\sigma_A\sigma_B}\right)\sigma_B \\ =\;& (\sigma_A - \sigma_B)\,\phi\left(\frac{1}{2}\frac{\sigma_B+\sigma_A}{\sigma_A\sigma_B}\right) < 0 \end{aligned}$$

This proves the proposition. ■

Proposition 4 *Suppose that the employer is "selective" in its hiring policy, i.e.,* $\underline{q} > p$*, then:*

1. As the employer becomes more selective, minority representation in the workplace decreases. Formally, r_B *is decreasing in* $\underline{q}$*.*

2. As the employer becomes arbitrarily selective, minorities vanish from the workplace. Formally, $\lim_{\underline{q}\to 1} r_B = 0$*.*

Proof. To prove part 1, differentiate r_B with respect to $\underline{q}$:

$$\begin{aligned}
\frac{\partial r_B}{\partial \underline{q}} &= \frac{-m_B g_{B1}\left(1 - m_B G_{B1} - m_A G_{A1}\right) - \left(-m_B g_{B1} - m_A g_{A1}\right) m_B \left(1 - G_{B1}\right)}{\left(1 - m_B G_{B1} - m_A G_{A1}\right)^2} \\
&= \frac{m_B m_A \left(g_{A1}\left(1 - G_{B1}\right) - g_{B1}\left(1 - G_{A1}\right)\right)}{\left(1 - m_B G_{B1} - m_A G_{A1}\right)^2}
\end{aligned}$$

Notice that the sign of $\frac{\partial r_B}{\partial \underline{q}}$ depends only on the hazard rates of G_{A1} and G_{B1}. And by Lemma 5 it then follows that $\frac{\partial r_B}{\partial \underline{q}} < 0$.

To prove part 2 of the proposition, notice that (via L'Hôpital's rule)

$$\lim_{\underline{q}\to 1} r_B = \lim_{\underline{q}\to 1} \frac{m_B}{m_B + m_A \frac{g_{A1}}{g_{B1}}}$$

and this limit depends solely on the limit of the likelihood ratio, $\frac{g_{A1}}{g_{B1}}$. Finally, it may be readily shown that:

$$\begin{aligned}
\lim_{\underline{q}\to 1} \frac{g_{A1}}{g_{B1}} &= \lim_{\underline{q}\to 1} \frac{\phi\left(\frac{s_A(q)-1}{\sigma_A}\right)\sigma_A}{\phi\left(\frac{s_B(q)-1}{\sigma_B}\right)\sigma_B} \\
&= \lim_{\underline{q}\to 1} e^{\frac{1}{8\sigma_A^2\sigma_B^2}\left(4\sigma_A^2\sigma_B^2 \ln^2\left(\frac{q}{1-q}\right)-1\right)\left(\sigma_B^2-\sigma_A^2\right)} \frac{\sigma_A}{\sigma_B} \to \infty
\end{aligned}$$

Hence,

$$\lim_{\underline{q}\to 1} r_B = 0$$

■

Proposition 5

1. Minorities are overrepresented among initial hires (i.e., $\frac{h_B}{m_B} > 1$*) iff* $0 < \underline{q}^* < \underline{q}^{\emptyset}$*.*

2. Minorities are underrepresented (i.e., $\frac{h_B}{m_B} \leq 1$*) iff* $\underline{q}^{\emptyset} < \underline{q}^* < 1$*.*

Proof. Argument identical to that in Proposition 3. ■

Proposition 6 *For all $\underline{q} \in (0,1)$, minority hires are fired at a higher rate than majority hires.*

Proof. Because hires are fired if and only if they turn out to be incompetent, we have to prove that

$$\Pr\left(\theta_A = 0 | q_A \geq \underline{q}\right) = \frac{(1-G_{A0})(1-p)}{1-G_A} < \frac{(1-G_{B0})(1-p)}{1-G_B} = \Pr\left(\theta_B = 0 | q_B \geq \underline{q}\right)$$

for all $\underline{q} \in (0,1)$.

This is equivalent to showing that

$$\frac{1-G_{A0}}{1-G_A} < \frac{1-G_{B0}}{1-G_B}$$

or

$$\frac{1-G_B}{1-G_{B0}} < \frac{1-G_A}{1-G_{A0}}$$

Now,

$$\begin{aligned}
\frac{1-G_B}{1-G_{B0}} &< \frac{1-G_A}{1-G_{A0}} \Longleftrightarrow \\
\frac{1-pG_{B1}-(1-p)\,G_{B0}}{1-G_{B0}} &< \frac{(1-pG_{A1}-(1-p)\,G_{A0})}{1-G_{A0}} \Longleftrightarrow \\
\frac{1-G_{B1}}{1-G_{B0}} &< \frac{1-G_{A1}}{1-G_{A0}}
\end{aligned}$$

Hence, showing that $\Pr\left(\theta_A = 0 | q_A \geq \underline{q}\right) < \Pr\left(\theta_B = 0 | q_B \geq \underline{q}\right)$ is equivalent to showing that the ratio of good hiring decisions over bad hiring decisions, $\frac{1-G_{\gamma 1}}{1-G_{\gamma 0}}$, is greater for kind A hires than for kind B hires. To prove the latter, we show that

$$\frac{d}{d\sigma_\gamma}\left[\frac{1-G_{\gamma 1}\left(\underline{q}\right)}{1-G_{\gamma 0}\left(\underline{q}\right)}\right] < 0$$

Now, $\frac{d}{d\sigma_\gamma}\left[\frac{1-G_{\gamma 1}\left(\underline{q}\right)}{1-G_{\gamma 0}\left(\underline{q}\right)}\right]$

$$\begin{aligned}
&= \frac{d}{d\sigma_\gamma}\left[\frac{\int_{\underline{q}}^1 g_{\gamma 1}(q)\,dq}{\int_{\underline{q}}^1 g_{\gamma 0}(q)\,dq}\right] \\
&= \frac{d}{d\sigma_\gamma}\left[\frac{\int_{\underline{q}}^1 \phi\left(\frac{s_\gamma(q)-1}{\sigma_\gamma}\right)\frac{\sigma_\gamma}{q(1-q)}dq}{\int_{\underline{q}}^1 \phi\left(\frac{s_\gamma(q)}{\sigma_\gamma}\right)\frac{\sigma_\gamma}{q(1-q)}dq}\right]
\end{aligned}$$

Using that $\frac{ds_\gamma(q)}{d\sigma_\gamma} = \frac{2\left(s_\gamma(q)-\frac{1}{2}\right)}{\sigma_\gamma}$, straightforward algebra leads to the conclusion that the sign of $\frac{d}{d\sigma_\gamma}\left[\frac{1-G_{\gamma 1}\left(\underline{q}\right)}{1-G_{\gamma 0}\left(\underline{q}\right)}\right]$ is equal to the sign of

$$\int_{\underline{q}}^{1} g_{\gamma 1}(q)\, dq \int_{\underline{q}}^{1} s_\gamma(q)\left(s_\gamma(q)-1\right) g_{\gamma 0}(q)\, dq - \int_{\underline{q}}^{1} g_{\gamma 0}(q)\, dq \int_{\underline{q}}^{1} s_\gamma(q)\left(s_\gamma(q)-1\right) g_{\gamma 1}(q)\, dq$$

Changing variables of integration from q to s, we get

$$\int_{s_\gamma\left(\underline{q}\right)}^{1} g_{\gamma 1}(s)\frac{\partial q_\gamma(s)}{\partial s} ds \int_{s_\gamma\left(\underline{q}\right)}^{1} s(s-1)\, g_{\gamma 0}(s)\frac{\partial q_\gamma(s)}{\partial s} ds$$
$$-\int_{s_\gamma\left(\underline{q}\right)}^{1} g_{\gamma 0}(s)\frac{\partial q_\gamma(s)}{\partial s} ds \int_{s_\gamma\left(\underline{q}\right)}^{1} s(s-1)\, g_{\gamma 1}(s)\frac{\partial q_\gamma(s)}{\partial s} ds$$

Substituting for $g_{\gamma 0}$, $g_{\gamma 1}$, and $\frac{\partial q_\gamma(s)}{\partial s}$,

$$\int_{s_\gamma\left(\underline{q}\right)}^{1} \phi\left(\frac{s-1}{\sigma_\gamma}\right) ds \int_{s_\gamma\left(\underline{q}\right)}^{1} s(s-1)\,\phi\left(\frac{s}{\sigma_\gamma}\right) ds - \int_{s_\gamma\left(\underline{q}\right)}^{1} \phi\left(\frac{s}{\sigma_\gamma}\right) ds \int_{s_\gamma\left(\underline{q}\right)}^{1} s(s-1)\,\phi\left(\frac{s-1}{\sigma_\gamma}\right) ds$$

Expanding $s(s-1)$,

$$\int_{s_\gamma\left(\underline{q}\right)}^{1} \phi\left(\frac{s-1}{\sigma_\gamma}\right) ds \left(\int_{s_\gamma\left(\underline{q}\right)}^{1} s^2\phi\left(\frac{s}{\sigma_\gamma}\right) ds - \int_{s_\gamma\left(\underline{q}\right)}^{1} s\phi\left(\frac{s}{\sigma_\gamma}\right) ds\right)$$
$$-\int_{s_\gamma\left(\underline{q}\right)}^{1} \phi\left(\frac{s}{\sigma_\gamma}\right) ds \left(\int_{s_\gamma\left(\underline{q}\right)}^{1} s^2\phi\left(\frac{s-1}{\sigma_\gamma}\right) ds - \int_{s_\gamma\left(\underline{q}\right)}^{1} s\phi\left(\frac{s-1}{\sigma_\gamma}\right) ds\right)$$

Writing in terms of conditional expectations,

$$\left(1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)\right)\left(1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)\right)\left(E\left[S_{\gamma 0}^2|S_{\gamma 0}\geq s_\gamma\left(\underline{q}\right)\right]-E\left[S_{\gamma 0}|S_{\gamma 0}\geq s_\gamma\left(\underline{q}\right)\right]\right)$$
$$-\left(1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)\right)\left(1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)\right)\left(E\left[S_{\gamma 1}^2|S_{\gamma 1}\geq s_\gamma\left(\underline{q}\right)\right]-E\left[S_{\gamma 1}|S_{\gamma 1}\geq s_\gamma\left(\underline{q}\right)\right]\right)$$

Dividing by the common positive factor $\left(1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)\right)\left(1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)\right)$:

$$E\left[S_{\gamma 0}^2|S_{\gamma 0}\geq s_\gamma\left(\underline{q}\right)\right]-E\left[S_{\gamma 0}|S_{\gamma 0}\geq s_\gamma\left(\underline{q}\right)\right]-E\left[S_{\gamma 1}^2|S_{\gamma 1}\geq s_\gamma\left(\underline{q}\right)\right]-E\left[S_{\gamma 1}|S_{\gamma 1}\geq s_\gamma\left(\underline{q}\right)\right]$$

Now, the moment generating function, mgf, of a left-truncated standard normal random variable U with truncation point d is (see, for example, Heckman and Honoré, 1990):

$$mgf(\beta) = e^{\frac{1}{2}\beta^2}\frac{\int_{d-\beta}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)du}{\int_{d}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)du}$$

Hence,

$$
\begin{aligned}
E\left[U|U \geq d\right] &= \frac{\partial mgf}{\partial \beta}|_{\beta=0} \\
&= \frac{\phi\left(d\right)}{1-\Phi\left(d\right)}
\end{aligned}
$$

while

$$
\begin{aligned}
E\left[U^2|U \geq d\right] &= \frac{\partial^2 mgf}{\partial \beta^2}|_{\beta=0} \\
&= 1 + d\frac{\partial mgf}{\partial \beta}|_{\beta=0} \\
&= 1 + \frac{d\phi\left(d\right)}{1-\Phi\left(d\right)}
\end{aligned}
$$

For $X \sim N\left(\mu, \sigma^2\right)$, this implies

$$
\begin{aligned}
E\left[X|X \geq d'\right] &= \mu + \frac{\sigma\phi\left(\frac{d'-\mu}{\sigma}\right)}{1-\Phi\left(\frac{d'-\mu}{\sigma}\right)} \\
E\left[X^2|X \geq d'\right] &= \sigma^2 + \left(\mu + d'\right)\frac{\sigma\phi\left(\frac{d'-\mu}{\sigma}\right)}{1-\Phi\left(\frac{d'-\mu}{\sigma}\right)} + \mu^2
\end{aligned}
$$

Now, recall that $S_{\gamma 0} \sim N\left(0, \sigma_\gamma\right)$ and $S_{\gamma 1} \sim N\left(1, \sigma_\gamma\right)$. Hence,

$$
E\left[S_{\gamma 0}^2|S_{\gamma 0} \geq s_\gamma\left(\underline{q}\right)\right] - E\left[S_{\gamma 0}|S_{\gamma 0} \geq s_\gamma\left(\underline{q}\right)\right] - E\left[S_{\gamma 1}^2|S_{\gamma 1} \geq s_\gamma\left(\underline{q}\right)\right] - E\left[S_{\gamma 1}|S_{\gamma 1} \geq s_\gamma\left(\underline{q}\right)\right]
$$

$$
\begin{aligned}
&= \sigma_\gamma^2 + s_\gamma\left(\underline{q}\right)\frac{\sigma_\gamma\phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)}{1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)} - \frac{\sigma_\gamma\phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)}{1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)} \\
&\quad - \left(\sigma_\gamma^2 + \left(1 + s_\gamma\left(\underline{q}\right)\right)\frac{\sigma_\gamma\phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)}{1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)} + 1 - 1 - \frac{\sigma_\gamma\phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)}{1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)}\right)
\end{aligned}
$$

Dividing by σ_γ and collecting terms, we get

$$
\left(s_\gamma\left(\underline{q}\right) - 1\right)\frac{\phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)}{1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)}{\sigma_\gamma}\right)} - s_\gamma\left(\underline{q}\right)\frac{\phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)}{1-\Phi\left(\frac{s_\gamma\left(\underline{q}\right)-1}{\sigma_\gamma}\right)}
$$

Hence, the question is whether

$$(s-1)\frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}-s\frac{\phi\left(\frac{s-1}{\sigma}\right)}{1-\Phi\left(\frac{s-1}{\sigma}\right)}<0$$

$$\frac{s-1}{\sigma}\frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}-\frac{s}{\sigma}\frac{\phi\left(\frac{s-1}{\sigma}\right)}{1-\Phi\left(\frac{s-1}{\sigma}\right)}<0$$

for all $s\in\mathbb{R}$ and $\sigma>0$.

Denote hazard rate $\frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}$ by $\lambda\left(\frac{s}{\sigma}\right)$. The expression then becomes

$$(s-1)\lambda\left(\frac{s}{\sigma}\right)-s\lambda\left(\frac{s-1}{\sigma}\right)$$

Graphically, when $s-1<0$, [**Figure 3 Here**]

Hence, for all $s-1<0$, it is obvious that

$$(s-1)\lambda\left(\frac{s}{\sigma}\right)-s\lambda\left(\frac{s-1}{\sigma}\right)<0$$

When $s-1>0$, graphically, [**Figure 4 Here**].

Here, in principle, it could go either way.

Now, for $s-1>0$,

$$\begin{aligned}
& (s-1)\lambda\left(\frac{s}{\sigma}\right)-s\lambda\left(\frac{s-1}{\sigma}\right)\\
= \; & (s-1)\left(\lambda\left(\frac{s}{\sigma}\right)-\lambda\left(\frac{s-1}{\sigma}\right)\right)-(s-(s-1))\lambda\left(\frac{s-1}{\sigma}\right)\\
\leq \; & \int_{\lambda\left(\frac{s-1}{\sigma}\right)}^{\lambda\left(\frac{s}{\sigma}\right)}\lambda^{-1}(l)\,dl-\int_{s-1}^{s}\lambda\left(\frac{x}{\sigma}\right)dx
\end{aligned}$$

where the inequality follows from the convexity of $\lambda\left(\frac{s}{\sigma}\right)$.

Changing the variable of integration in the first term from hazard rate l to signal x, the last expression becomes

$$\begin{aligned}
&= \int_{s-1}^{s}x\frac{\partial l}{\partial x}dx-\int_{s-1}^{s}\lambda\left(\frac{x}{\sigma}\right)dx\\
&= \int_{s-1}^{s}\frac{x}{\sigma}\lambda'\left(\frac{x}{\sigma}\right)dx-\int_{s-1}^{s}\lambda\left(\frac{x}{\sigma}\right)dx\\
&= \int_{s-1}^{s}\left(\frac{x}{\sigma}\lambda'\left(\frac{x}{\sigma}\right)-\lambda\left(\frac{x}{\sigma}\right)\right)dx
\end{aligned}$$

Finally, we show that the integrand, which we write as

$$z\lambda'(z)-\lambda(z)$$

is negative for all $z \geq 0$.

First, note that

$$\begin{aligned}
\lambda'\left(\frac{s}{\sigma}\right) &= \frac{d}{d\frac{s}{\sigma}}\lambda\left(\frac{s}{\sigma}\right)=\frac{d}{d\frac{s}{\sigma}}\left[\frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}\right] \\
&= \frac{-\frac{s}{\sigma}\phi\left(\frac{s}{\sigma}\right)\left(1-\Phi\left(\frac{s}{\sigma}\right)\right)+\phi^2\left(\frac{s}{\sigma}\right)}{\left(1-\Phi\left(\frac{s}{\sigma}\right)\right)^2} \\
&= \frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}\left(\frac{\phi\left(\frac{s}{\sigma}\right)-\frac{s}{\sigma}\left(1-\Phi\left(\frac{s}{\sigma}\right)\right)}{\left(1-\Phi\left(\frac{s}{\sigma}\right)\right)}\right) \\
&= \frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}\left(\frac{\phi\left(\frac{s}{\sigma}\right)}{1-\Phi\left(\frac{s}{\sigma}\right)}-\frac{s}{\sigma}\right) \\
&= \lambda\left(\frac{s}{\sigma}\right)\left(\lambda\left(\frac{s}{\sigma}\right)-\frac{s}{\sigma}\right)
\end{aligned}$$

Hence, the integrand can be written as

$$\begin{aligned}
& z\lambda'(z)-\lambda(z) \\
=\ & z\lambda(z)(\lambda(z)-z)-\lambda(z) \\
=\ & \lambda(z)(z(\lambda(z)-z)-1)
\end{aligned}$$

Dividing by $\lambda(z)$, The question becomes whether

$$z(\lambda(z)-z)<1$$

for $z \geq 0$.

Now, note that $\lambda'(z)<1$ for all z, as the derivative of the hazard rate of the standard Normal distribution converges to 1 from below when $z \to \infty$. Hence, it suffices to show that

$$z(\lambda(z)-z) \leq \lambda(z)(\lambda(z)-z)=\lambda'(z)$$

Now,

$$z(\lambda(z)-z) \leq \lambda(z)(\lambda(z)-z)$$

is equivalent to

$$0 \leq (\lambda(x)-x)^2$$

where the last inequality is obviously true. ■

C. Proofs of Implications

Implication 1 *An increase in the cost of firing, c, reduces workplace diversity.*

Proof. To establish part 1 of the implication, we show that $\underline{q}^*$ is increasing in c. Recall that optimality of the threshold strategy implies that

$$(V(\underline{q}^*) - v)\, \underline{q}^* + (1 - \underline{q}^*)\, c = 0 \tag{3}$$

Implicitly differentiating with respect to c while noting that $\frac{\partial V(\underline{q}^*)}{\partial \underline{q}^*} = 0$ gives

$$(V(\underline{q}^*) - v)\frac{d\underline{q}^*}{dc} + \frac{\partial V(\underline{q}^*)}{\partial c}\underline{q}^* + (1 - \underline{q}^*) - c\frac{d\underline{q}^*}{dc} = 0$$

Solving for $\frac{d\underline{q}^*}{dc}$:

$$\frac{d\underline{q}^*}{dc} = \frac{\left(\frac{\partial V(\underline{q}^*)}{\partial c} - 1\right)\underline{q}^* + 1}{v + c - V(\underline{q}^*)}$$

It is easily checked that

$$\frac{\partial V(\underline{q}^*)}{\partial c} = \frac{-\delta \int_{\underline{q}^*}^1 (1-q)\, dG(q)}{1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)}$$

Substituting into the expression for $\frac{d\underline{q}^*}{dc}$ and simplifying, one obtains

$$\frac{d\underline{q}^*}{dc} = \left(\frac{\delta G(\underline{q}^*) - 1}{1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)}\right)\underline{q}^* + 1$$

To establish that the right-hand side of this expression is positive requires that we show

$$(1 - \delta G(\underline{q}^*))\, \underline{q}^* - \left(1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)\right) < 0$$

To see this, notice that

$$\begin{aligned}
& (1 - \delta G(\underline{q}^*))\, \underline{q}^* - \left(1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)\right) \\
< \; & (1 - \delta G(\underline{q}^*))\, \underline{q}^* - \left(1 - \delta\left(1 - \underline{q}^*\left(1 - G(\underline{q}^*)\right)\right)\right) \\
= \; & -(1-\delta)(1 - \underline{q}^*) \\
< \; & 0
\end{aligned}$$

■

Implication 2 *A decrease in the cost of interviewing, k, reduces workplace diversity.*

Proof. To establish the implication, we show that $\underline{q}^*$ is decreasing in k. Implicitly differentiating equation (3) with respect to k while noting that $\frac{\partial V(\underline{q}^*)}{\partial \underline{q}^*} = 0$, we obtain

$$\left(V\left(\underline{q}^*\right) - v\right)\frac{d\underline{q}^*}{dk} + \frac{\partial V\left(\underline{q}^*\right)}{\partial k}\underline{q}^* - c\frac{d\underline{q}^*}{dk} = 0$$

Solving for $\frac{d\underline{q}^*}{dk}$,

$$\frac{d\underline{q}^*}{dk} = \frac{\frac{\partial V\left(\underline{q}^*\right)}{\partial k}\underline{q}^*}{v + c - V\left(\underline{q}^*\right)}$$

Hence, $\frac{d\underline{q}^*}{dk}$ and $\frac{\partial V(\underline{q}^*)}{\partial k}$ have the same sign, while it is easily checked that $\frac{\partial V(\underline{q}^*)}{\partial k} < 0$. ■

Implication 3 *Diversity is procyclical. Formally, q^* is decreasing in v (and k).*

Proof. From Implication 1, we already know that q^* is increasing in k.

Implicitly differentiating equation (3) with respect to v while noting that $\frac{\partial V(\underline{q}^*)}{\partial \underline{q}^*} = 0$, we obtain

$$\left(V\left(\underline{q}^*\right) - v\right)\frac{d\underline{q}^*}{dv} + \left(\frac{\partial V\left(\underline{q}^*\right)}{\partial v} - 1\right)\underline{q}^* - c\frac{d\underline{q}^*}{dv} = 0$$

Solving for $\frac{d\underline{q}^*}{dv}$:

$$\frac{d\underline{q}^*}{dv} = \frac{\left(\frac{\partial V\left(\underline{q}^*\right)}{\partial v} - 1\right)\underline{q}^*}{v + c - V\left(\underline{q}^*\right)}$$

It is easily checked that

$$\frac{dV\left(\underline{q}^*\right)}{dv} = \frac{\delta \int_{\underline{q}^*}^1 (q)\, dG(q)}{1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)}$$

Substituting this back into $\frac{d\underline{q}^*}{dv}$ and simplifying, one obtains

$$\begin{aligned}\frac{d\underline{q}^*}{dv} &= -\frac{\frac{1-\delta}{1-\delta\left(1-\int_{\underline{q}^*}^1 q dG(q)\right)}\underline{q}^*}{v + c - V\left(\underline{q}^*\right)} \\ &< 0\end{aligned}$$

■

Implication 4 *Riskier firms are more diverse. Formally, $\underline{q}^*$ is increasing in δ.*

Proof. Implicitly differentiating equation (3) with respect to δ while noting that $\frac{\partial V\left(\underline{q}^*\right)}{\partial \underline{q}^*} = 0$, we obtain

$$\left(V\left(\underline{q}^*\right) - v\right)\frac{d\underline{q}^*}{d\delta} + \left(\frac{dV\left(\underline{q}^*\right)}{d\delta} - 1\right)\underline{q}^* - c\frac{d\underline{q}^*}{d\delta} = 0$$

Solving for $\frac{d\underline{q}^*}{d\delta}$:

$$\frac{d\underline{q}^*}{d\delta} = \frac{\left(\frac{dV\left(\underline{q}^*\right)}{d\delta} - 1\right)\underline{q}^*}{v + c - V\left(\underline{q}^*\right)}$$

It is easily checked that:

$$\frac{dV\left(\underline{q}^*\right)}{d\delta} = \frac{Z\left(1 - \delta X\right) + X\left(\delta Z - k\right)}{\left(1 - \delta X\right)^2}$$

where

$$\begin{aligned} Z &\equiv \int_{\underline{q}}^{1} \left(qv + \left(1 - q\right)\left(-c\right)\right) dG\left(q\right) \\ X &\equiv \left(1 - \int_{\underline{q}}^{1} q dG\left(q\right)\right) \end{aligned}$$

To show that $\frac{d\underline{q}^*}{d\delta} > 0$, it is sufficient to show that $\frac{dV\left(\underline{q}^*\right)}{d\delta} - 1 > 0$, or equivalently

$$Z\left(1 - \delta X\right) + X\left(\delta Z - k\right) - \left(1 - \delta X\right)^2 > 0$$

To see this, simplify the left-hand side of the above expression and recall that, since the employer finds it optimal to search in the first place, $\delta Z - k \geq 0$. This yields

$$\begin{aligned} & Z - Xk + \left(1 - X\delta\right)^2 \\ \geq\ & Z - X\delta Z + \left(1 - X\delta\right)^2 \\ =\ & \left(1 - X\delta\right)\left(Z + 1 - X\delta\right) \\ >\ & 0 \end{aligned}$$

where the last inequality follows from the fact that $Z > 0$ and X, $\delta \in (0, 1)$. ■

Implication 5 *In jobs that require exceptional skills, minorities will be underrepresented. In jobs that require common skills, minorities will be overrepresented. Formally, there exists* $0 < p_0 < p_1 < 1$ *such that, for all* $p \in (0, p_0)$, $\frac{r_B}{m_B} < 1$, *while for all* $p \in (p_1, 1)$, $\frac{r_B}{m_B} > 1$.

Proof. First, we establish that $\lim_{p\uparrow 1} \underline{q}^* < 1$ and $\lim_{p\downarrow 0} \underline{q}^* > 0$. To see this, note that $\underline{q}^*$ is monotone in p since, by implicitly differentiating equation (3),

$$\frac{d\underline{q}^*}{dp} = \frac{\frac{\partial V(\underline{q}^*)}{\partial p}\underline{q}^*}{v - V\left(\underline{q}^*\right) + c} > 0$$

where the inequality follows from the fact that $v > V(\underline{q}^*)$ and, by Lemma 1, $\frac{\partial V(\underline{q}^*)}{\partial p} > 0$.

Since $\underline{q}^*$ is bounded and monotone function of p we know that both limits must exist.

To establish that $\lim_{p\uparrow 1} \underline{q}^* < 1$, suppose, to the contrary, that $\lim_{p\uparrow 1} \underline{q}^* = 1$. Then the right-hand side of equation (2) becomes:

$$\begin{aligned} & \lim_{p\uparrow 1} \frac{\left(1 - \delta\left(1 - \int_1^1 q dG(q)\right)\right)c}{(1 - \delta G(1))\, c + (1 - \delta)\, v + k} \\ = & \frac{(1-\delta)\, c}{(1-\delta)\, c + (1-\delta)\, v + k} \neq 1 \end{aligned}$$

which is a contradiction.

To establish that $\lim_{p\downarrow 0} \underline{q}^* > 0$, recall that $\underline{q}^*$ is implicitly defined by equation (2). Taking limits:

$$\begin{aligned} \lim_{p\downarrow 0} \underline{q}^* &= \lim_{p\downarrow 0} \frac{\left(1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)\right)c}{\left(1 - \delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\, v + k} \\ &> \lim_{p\downarrow 0} \frac{(1-\delta)\, c}{c + (1-\delta)\, v + k} > 0 \end{aligned}$$

To complete the proof, it remains to show that $\underline{q}^0$ and $\underline{q}^1$ are monotone in p with limits $\lim_{p\downarrow 0} \underline{q}^0 = 0$ and $\lim_{p\downarrow 1} \underline{q}^1 = 1$. Monotonicity may be readily verified by differentiating the expressions for $\underline{q}^0$ and $\underline{q}^1$. Likewise, the limit results are trivial to obtain. ■

Implication 6 *The larger the minority fraction of the population, the smaller its degree of underrepresentation in the workforce. Formally,* $\underline{q}^*$ *is decreasing in* m_B.

Proof. Recall that $\underline{q}^*$ satisfies

$$\begin{aligned} \underline{q}^* &= \frac{\left(1 - \delta\left(1 - \int_{\underline{q}^*}^1 q dG(q)\right)\right)c}{\left(1 - \delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\, v + k} \\ &= \frac{(1-\delta)\, c + c\delta \int_{\underline{q}^*}^1 q dG(q)}{\left(1 - \delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\, v + k} \end{aligned}$$

$$= \frac{(1-\delta)\,c + c\delta\left(\int_{\underline{q}^*}^{1} qdG(q) + \int_0^{\underline{q}^*} qdG(q) - \int_0^{\underline{q}^*} qdG(q)\right)}{\left(1-\delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\,v + k}$$

$$= \frac{(1-\delta)\,c + c\delta\left(E_G[Q] - \int_0^{\underline{q}^*} qdG(q)\right)}{\left(1-\delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\,v + k}$$

$$= \frac{(1-\delta)\,c + c\delta\left(E_G[Q] - \left(\underline{q}^* G\left(\underline{q}^*\right) - \int_0^{\underline{q}^*} G(q)\,dq\right)\right)}{\left(1-\delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\,v + k}$$

$$= \frac{(1-\delta)\,c + c\delta\left(E_G[Q] - \underline{q}^* G\left(\underline{q}^*\right) + \int_0^{\underline{q}^*} G(q)\,dq\right)}{\left(1-\delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\,v + k}$$

Therefore,

$$\begin{aligned}
\underline{q}^*\left(\left(1-\delta G\left(\underline{q}^*\right)\right)c + (1-\delta)\,v + k\right) &= (1-\delta)\,c + c\delta\left(E_G[Q] - \underline{q}^* G\left(\underline{q}^*\right) + \int_0^{\underline{q}^*} G(q)\,dq\right) \\
(c + (1-\delta)\,v + k)\,\underline{q}^* - c\delta\underline{q}^* G\left(\underline{q}^*\right) &= (1-\delta)\,c + c\delta E_G[Q] - c\delta\underline{q}^* G\left(\underline{q}^*\right) + c\delta\int_0^{\underline{q}^*} G(q)\,dq \\
(c + (1-\delta)\,v + k)\,\underline{q}^* &= (1-\delta)\,c + c\delta\left(E_G[Q] + \int_0^{\underline{q}^*} G(q)\,dq\right)
\end{aligned}$$

Now, from Lemma 2, we know that if $m_B < m_B'$, then $G(q, m_B)$ is a mean preserving spread of $G(q, m_B')$. Hence, if we go from m_B to m_B', $E_G[Q]$ remains unchanged in the RHS of the last equation but, by definition of second-order stochastic dominance, $\int_0^{\underline{q}^*} G(q; m_B) > \int_0^{\underline{q}^*} G(q; m_B')$. Hence, the LHS also increases. Therefore, it must be that $\underline{q}^*(m_B) > \underline{q}^*(m_B')$, because c, δ, v, and k are all constants. We conclude that $\frac{\partial \underline{q}^*}{\partial m_B} < 0$. ■

References

Aigner, Dennis, and Glen Cain. 1977. "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30: 175–87.

Arrow, Kenneth J. 1998. "What Has Economics to Say about Racial Discrimination," *Journal of Economic Perspectives*, 12(2): 91–100.

Athey, Susan, Christopher Avery and Peter Zemsky. 2000. "Mentoring and Diversity," *American Economic Review*, 90(4): 765–786.

Becker, Gary S. 1957, *The Economics of Discrimination.* Chicago: Chicago University Press.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 94(4): 991–1013.

Black, Dan A. 1995. "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics*, 13(2): 309–33.

Bolch, Matt. 2007. "Lights, Camera ... Interview!" *HR Magazine*, 52(3).

Coate, Stephen, and Glenn Loury. 1993. "Will Affirmative Action Eliminate Negative Stereotypes?"*American Economic Review*, 83(5): 1220–40.

Cornell, Bradford, and Ivo Welch. 1996. "Culture, Information and Screening Discrimination," *Journal of Political Economy*, 104(3): 542–571.

Dagevos, Jaco. 2006, *Hoge (jeugd)werkloosheid onder etnische minderheden. Nieuwe bevindingen uit het LAS-onderzoek.*" Den Haag: Sociaal en Cultureel Planbureau.

Heckman, James, and Bo Honoré. 1990. "The Empirical Content of the Roy Model," *Econometrica*, 58(5): 1121–1149.

Lang, Kevin. 1986. "A Language Theory of Discrimination, " *Quarterly Journal of Economics*, 101(2): 363-382.

McCall, John J. 1970. "Economics of Information and Job Search," *Quarterly Journal of Economics*, 84(1): 113–26.

Lundberg, Shelly J., and Richard Startz. 1998. "On the Persistence of Racial Inequality," *Journal of Labor Economics*, 16(2): 292–323.

Lundberg, Shelly J., and Richard Startz. Forthcoming. "Information and Racial Exclusion," *Journal of Population Economics.*

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62(4): 659–61.

Rosen, Asa. 1997. "An Equilibrium Search-Matching Model of Discrimination," *European Economic Review*, 41(8): 1589–1613.

Scollon, Ronald, and Suzanne Scollon. 2001. *Intercultural Communication: A Discourse Approach.* Boston: Blackwell Publishers.

Stigler, George. 1961. "The Economics of Information," *Journal of Political Economy*, 69(3): 213–25.

Tesser, Paul., Ans Merens and Carlo van Praag. 1999. *Rapportage minderheden 1999: positie in het onderwijs en op de arbeidsmarkt.* Den Haag: Sociaal en Cultureel Planbureau.

*Morgan: Department of Economics and Haas School of Business, University of California, Berkeley, Berkeley, CA 94720 USA, E-mail: morgan@haas.berkeley.edu.

Várdy: International Monetary Fund, 700 19th Street, NW, Washington, DC 20431. E-mail: fvardy@imf.org. Corresponding author.

We would like to thank two anonymous referees, Mary Amiti, Bob Feldman, Andrew Feltenstein, Harold Houba, Sunil Sharma, Keith Takeda and, especially, Johan Walden for extremely useful comments. The first author gratefully acknowledges the financial support of the National Science Foundation. The views expressed in this paper are those of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

Figure 1: Hiring probability ratios of competent candidates of kind A versus B : $\frac{1-G_{A1}(q)}{1-G_{B1}(q)}$.

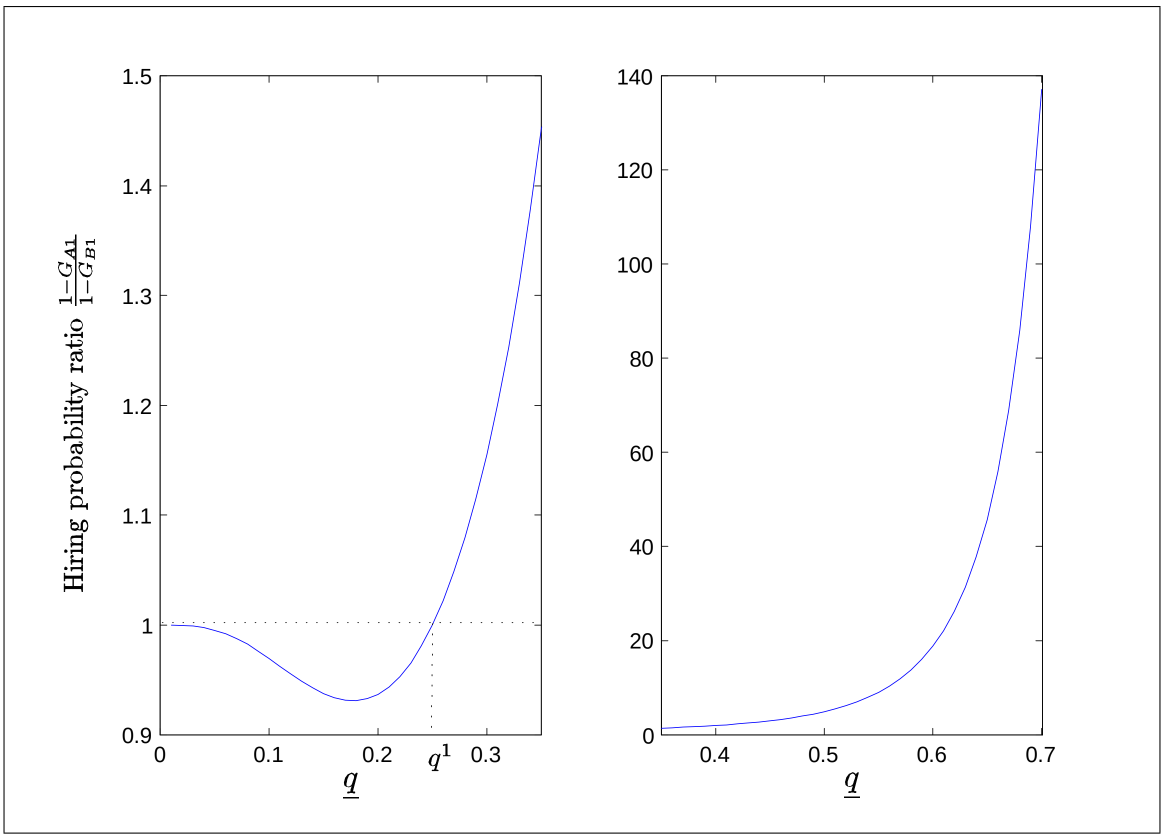

Figure 2: Over and Underrepresentation of Minorities.

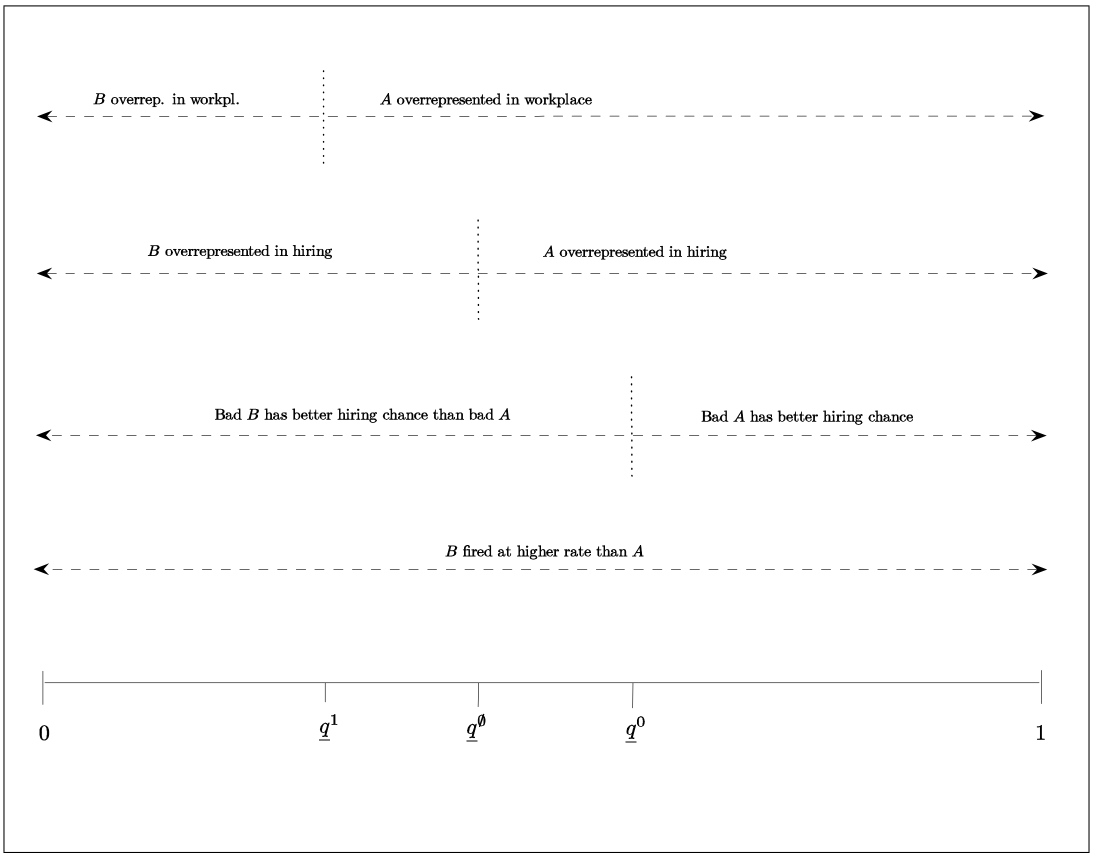

Figure 3:

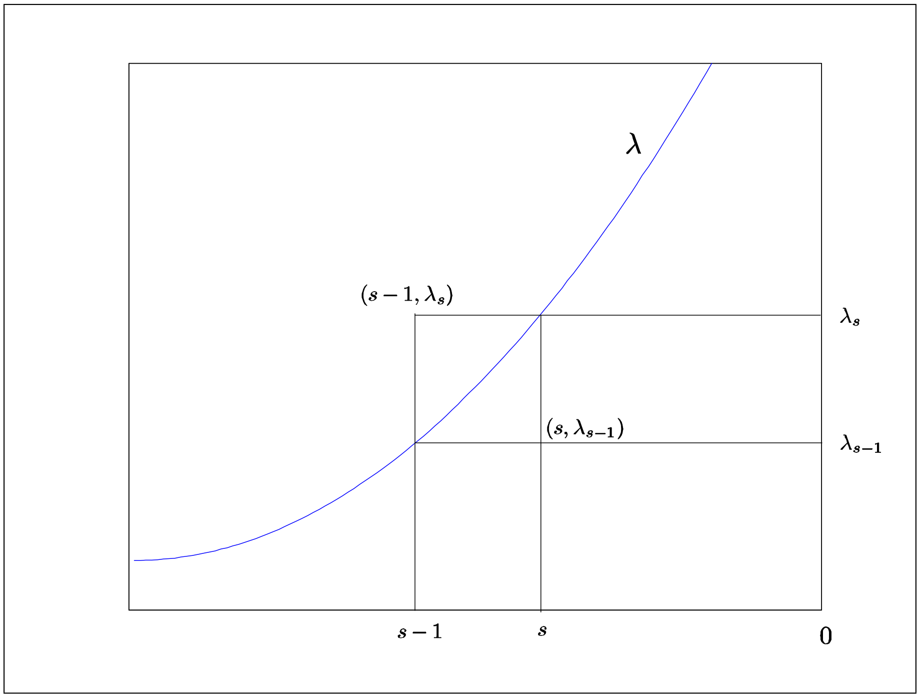

Figure 4: