INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
RICE UNIVERSITY

LEGAL, INFORMATIONAL, AND TESTING RESTRICTIONS IN EMPLOYMENT: A STUDY OF ANTIDISCRIMINATION POLICY

by

JOYCE CAROL SADKA-NEGRIN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE

Dagobert Brito, Chairman
Peterkin Professor of Political Economy

Peter Hartley
Professor of Economics

Allen Matusow
Professor of History

HOUSTON, TEXAS
APRIL 1998
Abstract

This dissertation addresses the problem of informational and legal restrictions on employment screening processes such as educational requirements and standardized testing. For the past two decades, the legal system in the U.S. has struggled with this issue and has labeled certain information or tests as discriminatory because they have adverse impact on minority groups. In the first chapter, we consider the major legal cases in this area and discuss the rationale for these major decisions in light of the evidence on testing and information from the fields of psychometrics and economics. We find that the legal system must take into account the productive inefficiencies that can be caused by informational restrictions, as well as the effects on incentives of firms and workers involved in these legal limitations. In chapter 2, we develop an extension of the language model of discrimination, adding differences in worker ability and the use of language knowledge as a signal of ability. We show that legal restrictions unambiguously lower output. They also result in lower wages for low ability minority workers, without increasing the wage for high ability minorities. Thus informational restrictions have an adverse and regressive effect on the minority income distribution. In chapter 3, we allow firms two selection devices, a standardized test and a language interview. We also consider two possible firm technologies. We find that firms will react to testing restrictions by producing through methods that are less sensitive to employee ability. This results in a regressive wage effect, even in the case of one language group. With two language groups, we find that integration tends to increase with restrictions on standardized testing, but production is reduced and low ability minority wages fall.
Acknowledgements

I would like to thank my advisor, Professor Dagobert Brito, for his sound advice and his patience and faith in me during these past four years. I would also like to thank Professor Peter Hartley for many hours of help with modeling and intuition. I thank Professor Allen Matusow for serving on my committee and for his insightful and positive questions. Professor Peter Mieszkowski first introduced me to the topic of my dissertation and I thank him for getting me started with this project. I benefited from discussions with Professor Kei-Mu Yi and Professor Marc Dudey. I would like to thank all my professors at Rice for the training they provided me with, and I thank the secretaries at the department for kindness and help. I would like to thank Joy Bryant especially for being a friend and support to graduate students in our department.

I believe that I could not have made it through graduate school without a strong network of moral support that was provided to me by many good friends. I cannot name all of them, but I would like to thank some friends especially: my classmates and colleagues Jenny Williams, Sly Gaudin, Jan van Heerden, Esperanza Paredes, Michelle Packer, Craig Johnson, Juan Rosellon, Joya Mukherjee, and Deba Ghosh. I also want to thank Guillermo Rabiela for his support and the many discussions we have had about economics. I thank Analia and Pablo Girault and Jill Garde for their much-needed encouragement over the years. I want to thank Rev. and Mrs. Richard Butts for their friendship and help and hospitality with a foreign student. I would like to thank my
friends and Robert Thomson and Thomas Page for their support and for making life much more fun.

I would like to thank my sister Nancy for being there for me at many important junctures of graduate school. I thank my brother Paul and his wife Nancy for their support and for many good vacations in their home. I want to thank my brother Richard, for so many things: for getting me interested in economics to start with, for all his help in the early years of classes at Rice, for his encouragement and help and kindness during this dissertation process. I want to thank my friend Nelly, for praying and caring and writing to encourage me when graduate life was not going too well. I want to thank my parents. I owe to them so many of the resources that have brought me through graduate school. I thank them for their vision for me, for their investment in me, for their prayers and support and visits. I want to thank them for always wanting the best for me no matter where that took me, and for giving up so much in order to give me the best education possible. I want to thank for motivating me to enjoy learning, and for the confidence they instilled in me so early in life, that there was nothing good and laudable in education that I could not accomplish.

I would like to dedicate this work to my husband, Jose Luis. I do not have any words to describe my thanks for all his support and help for the past 7 years. I am sure I would not be writing these acknowledgements were it not for him; I would not be who I am without him. I thank him for his faith in me, for countless and continuous discussions about economics and hundreds of other things, for being there with me at all the worst and best
times of this process. Thank you esposo. I would like to thank God for all those he has sent my way to help and guide me, like all those mentioned here. I want to thank Him for the abilities He has given me to do this work. Every good gift is from Him, and I thank Him with all my heart.
# Table of Contents

Chapter 1: Three Perspectives on Employment Testing  
1.1 Introduction  
1.2 The Legal Perspective  
1.2.1 Plaintiff's Prima Facie Case  
1.2.1.1 Disparate Treatment  
1.2.1.2 Griggs  
1.2.1.3 Fairness to groups vs. Fairness to Individuals  
1.2.1.4 The Relevant Application Pool  
1.2.1.5 Wards Cove and the Civil Rights Act of 1991  
1.2.1.6 Statistical Techniques and the EEOC Guidelines  
1.2.2 Employer Defenses  
1.2.2.1 Bona Fide Occupational Qualification  
1.2.2.2 Business Necessity and Job Relatedness  
1.2.3 Alternative Employment Practices  
1.3 The Psychometric Perspective  
1.3.1 The “Best Single Predictor” Evidence  
1.3.2 Alternative Tests  
1.4 The Economic Perspective  
1.4.1 The Basic Informational Problem  
1.4.2 The Basis of Discriminatory Outcomes  
1.4.3 Incentive Effects of Restrictions  
1.5 Conclusion  

Chapter 2: Language Discrimination, Signaling, and Informational Restrictions  
2.1 Introduction  
2.2 Models of Labor Market Discrimination  
2.3 Evidence in Favor of the Language Discrimination Model  
2.4 Adding Ability Differences to the Language Discrimination Model  
2.5 Effects of Legal Restrictions  
2.6 Conclusion  

Chapter 3: Standardized Testing Restrictions and Employment Discrimination  
3.1 Introduction  
3.2 Background and Motivation  
3.3 Model of Discrete Ability and Testing Outcomes  
3.3.1 Technologies  
3.3.2 Standardized Testing  
3.3.3 Benchmark Case: Perfect Information, One Speech Group  
3.3.4 Imperfect Information, One Speech Group  
3.3.5 Imperfect Information, Two Speech Groups  
3.3.6 Enforced Segregation  
3.3.7 Integration is Allowed  
3.4 Continuous Testing Outcomes, One Speech Group  
3.5 Conclusion  

Appendix 1: Discrete Case, $p = \frac{1}{4}$
CHAPTER 1: Three Perspectives on Employment Testing

1.1 Introduction

Among the employment practices that have come under close scrutiny since the 1964 Civil Rights Act are standardized testing procedures used by firms for hiring and promotion. Standardized tests have been used extensively by industrial firms; thus legal restrictions on such selection procedures can have important effects on firm behavior, production, and wages. In order to formulate coherent public policy in this area, some basic questions must be asked: Why do we want to restrict such tests? What are the benefits from imposing restrictions, and what are the losses? Knowing the tradeoff involved, it is worthwhile to implement this restriction policy?

Three disciplines are particularly concerned with the standardized testing issue. The legal profession creates the rules that restrict tests or loosen testing restrictions. Due to the deliberate generality of statutory language, the major debate over testing has taken place in courts and not in Congress. The legal system has originated a majority of the testing restrictions, and has elaborated on its policy motives for imposing these limitations on firms.

Psychometricians are primarily concerned with testing because they have created standardized tests and continue to refine them and study their characteristics. As such, psychometrics is very attentive to the legal system’s evaluation of the efficacy of
employment tests. As testing restrictions have increased, the level of psychometric sophistication in formulating and validating standardized tests has increased dramatically.

Finally, economists are interested in the testing issue because it concerns the labor market and represents a complex informational problem. A testing restriction is equivalent to increasing the cost of certain informational exchanges. Before any restrictions were imposed, the economist assumes that firms used standardized tests with their usual profit maximizing motive. Therefore testing must be in some way productive, and its elimination may cause losses in efficiency and production. It is also likely to affect all agents’ incentives, as well as the levels of certain activities.

We will argue that a proper understanding of the testing problem requires a synthesis of the different points of view of the above professions. In particular, no consistent legal or public policy towards testing is possible without taking account of the evidence and arguments from psychometrics and economics. Because legal policy must take account of the benefits of testing, psychometric evidence regarding the usefulness of tests must be examined. On the other hand, because agents involved in the employment testing controversy are likely to be rationally seeking their own best interest, legal rules must take account of productivity costs and effects on employer and worker incentives.

Section 1.2 presents the legal aspect of the testing issue. We follow the development of a typical employment discrimination and explain the relevance of testing evidence at each stage of the case. Section 1.3 discusses the psychometric evidence on testing and
alternative selection procedures. Section 1.4 presents the question from an economic perspective and describes the tradeoffs involved in the elimination of testing. Section 1.5 concludes and offers policy recommendations.

1.2 The legal perspective

The legal profession continues to struggle with employment testing, both at the general policy level, as discussed above, and at a more practical level. For example, courts puzzle over the practical question: What kinds of tests will pass the scrutiny of the Civil Rights Acts of 1964 and 1991? We know that a test which clearly a facially neutral selection criteria, can be found, without any evidence of intent, to constitute a violation of Title VII. In light of this law, employers want to know what they can do to ensure no liability in a Title VII suit, and to discourage plaintiffs from bringing such suits against them. Is it enough for employers to keep statistics and follow studies which show that their test is “job related” and does predict employee performance? Must they show that it is unbiased and equally accurate towards different groups of workers? Must they show that there is no other selection criteria which would have less of a disparate impact and would nevertheless provide the necessary “job related” information? If there is such an alternative selection procedure, does the disparity in cost, between that procedure and the allegedly discriminatory one, justify the use of the criterion with more adverse impact? Should employers stick to non-objective selection criteria and make sure their bottom line does not show disparate impact on any protected group, possibly through hiring quotas? To better understand these questions, we turn to the development of a typical discrimination test involving testing evidence.
1.2.1 Plaintiff’s prima facie case

The plaintiff must establish a prima facie case in order for the lawsuit to proceed. If the plaintiff is successful, the burden of rebutting the case shifts to the defendant. In other words, if a defendant fails to answer a successful prima facie argument, he loses the case.

1.2.1.1 Disparate treatment

The Supreme Court turned its attention to standardized employment testing at the same time that it accepted the disparate impact theory of employment discrimination, in Griggs1. Nevertheless, a prima facie case of disparate treatment can be shown on the basis of the employer’s use of a standardized test. The plaintiff must show that the employer’s test resulted in a pattern of worker selection or promotion that cannot be rationally explained without discriminatory intent. For example, a plaintiff could show that the racial composition of the employer’s work force selected by the test in question was less than 5% likely to have occurred by chance, given lack of discriminatory intent. The Court has argued that “absent explanation, it is ordinarily to be expected that nondiscriminatory hiring practices will in time result in a work force more or less representative of the racial composition of the general population”2. Thus, using statistical evidence of gross disparate impact on a protected Title VII group was allowed as proof of discriminatory intent before disparate impact was accepted as a theory of liability under the Civil Rights Act of 1964.

---

On the other hand, the Court has stated that more than purely statistical evidence is required for prima facie proof of intent\(^3\). Direct or indirect (circumstantial) proof of intent may include historical background and a marked departure of the employer from accepted practice within the industry, in addition to the adverse impact\(^4\). For instance, a plaintiff might succeed in showing that the employer's test is almost never used for screening in the industry, or that it is well-known to the employer to be a test biased against minority groups\(^5\). Such evidence, added to a clear showing of adverse impact, would strengthen the disparate treatment case significantly.

A disparate treatment showing is always desirable for the plaintiff. Courts are far more uniform in their dealings with adverse treatment cases, employer justifications are far more difficult (as we shall see under defenses), and damages are potentially much higher. Yet most of the standardized testing cases are tried under disparate impact, due to the difficulties of inferring intent from a facially neutral selection criterion.

1.2.1.2 Griggs

*Griggs v. Duke Power* dramatically altered the Title VII landscape. Although the Court upheld findings that Duke Power had no discriminatory intent in its screening procedures, the employer was nevertheless guilty of discrimination under Title VII. This finding was based on two elements. First, the screening procedures, consisting of employment tests

\(^5\) *Washington* and *City of Arlington Heights* were both equal protection cases against government employers or agencies. However, in *Washington* the Court found that proving a violation of equal protection was equivalent to proving disparate treatment under Title VII.
and educational requirements, were found to have a clear adverse impact on black workers. Second, the employer did not show that the screening procedures were “job-related and consistent with business necessity”. The Court discussed the language of Title VII §703(h), which allows professionally developed tests “not used to discriminate”. Referring to the EEOC guidelines, the Court interpreted this section as allowing tests which measure the knowledge or skills required for the performance of a specific job or set of jobs\(^6\).

The *Griggs* Court rejected the idea that the tests were not only used to hire employees with the minimum qualifications for entry level jobs, but also to obtain a work force capable of advancing to more complex job assignments. Pre-1965 employees who did not have to satisfy the new criteria for advancement were promoted at roughly the same rates as newer employees who did have to satisfy this criteria. Thus the Court reasoned that the new screening techniques were not even useful as a “long range” job qualification test.

This reasoning is flawed because the Court did not take into account possibly higher firing rates for workers who did not have to meet the new screening criteria. Imagine that after some period of employment, the firm can tell how productive a given worker is. Unproductive workers are fired and those who remain are promoted over time. Suppose also that there are two possible screening technologies, A and B. “A” is the older technology, resulting in 25% “bad apples” in the worker pool; B is an improved

technology, resulting in only 15% "bad apples". During the first year of employment, the bad workers are found out and fired. We consider the replacement of technology A by B one year (or more) after the change in policy. Since the bad workers from technology "A" days have already been weeded out, we will observe that promotion rates for those hired under A and B are similar. The Griggs Court would then infer that B did not add anything to the screening efficiency of the firm. This amounts to disregarding all the costs associated with one full year of training, monitoring, and paying the extra 10% bad workers that enter the worker pool under technology A. While this particular argument may not have worked in Griggs (because firing rates may have been available and similar between the two groups), the general assumption that similar promotion rates before and after a change in testing procedures implies the uselessness of the new procedure is false.

Griggs also established a judicial disfavor towards generalized aptitude testing. Although this was unnecessary for its holding, the Court chose to despise "broad or general testing devices" as "inadequate" for screening job applicants. It went on to support this assertion by maintaining that history is full of great achievers who did not have diplomas and credentials. This portion of the opinion is surprising for two reasons. First, history notes the achievements of "uncredentialled" heroes precisely because they are the exceptions to the rule. While this is good and well as something to admire and aspire to, the Court can hardly expect employers to gamble on a large percentage of their work force being uncredentialled over-achievers. Firms survive by taking advantage of such rules predicting worker performance, not by disregarding them. Second, as we shall see in the section on psychometrics, it is simply not true that general aptitude testing is inadequate
as a selection tool for many jobs. Even in 1971, there was ample evidence that such standardized tests were very good predictors of worker performance for a wide range of basic industrial jobs. After *Griggs*, this evidence only accumulated further. The test used in this case may have been biased against minorities or highly inaccurate in predicting their performance, as some tests were in 1971. But that would be a criticism of the construction of the test, not of its very use.

1.2.1.3 Fairness to groups vs. fairness to individuals

The anti-discrimination laws are based on the premise that, being alike in every relevant way, workers should not be treated differently on the basis of their group affiliation. The Court has discussed fairness to individuals as the crucial requirement placed on employers. In *Los Angeles Department of Water v. Manhart* (1978), the Court argued that using different actuarial tables of longevity for male and female employees was unfair to individual women because they could live far fewer years than the average woman. Yet, under Griggs, the fact that black workers nationwide failed the test at significantly higher rates than whites implied a lack of fairness to a set of black workers, because they had failed the test and belonged to the group of black workers nationwide; i.e. the test was unfair to black workers because it was unfair to their group. Perhaps the Court meant that *both* fairness to groups and individuals is mandated. Unfortunately, the two are often incompatible, as evidenced by *Manhart*. Requiring the same pension contributions from men and women would have resulted in male employees effectively subsidizing female employees as a group; this policy would clearly be unfair to men as a group. The tension between group and individual fairness is supposedly resolved by the
business necessity defense. Group impact of selection procedures is allowed to differ inasmuch as this difference is related to an underlying divergence in abilities or preferences between the groups. However, the likelihood of convincing a court of underlying ability differences without invoking past discrimination is very low. If group differences depend on past discrimination, which they often do, courts are likely to find the employer guilty, under the theory that the test used “perpetuates” a difference caused by past discrimination.

1.2.1.4 The relevant applicant pool

The plaintiff must show adverse impact of the selection procedure or test on the applicable pool of workers. This “applicable pool” has been the subject of some controversy. To begin with, although the applicable pool in the narrowest sense is the actual job applicant pool (or candidates for promotion), evidence from much larger samples has been allowed as proof of adverse impact7.

Such proof is allowed for two reasons. First, due to small sample statistical problems, the plaintiff may have very little chance of proving anything one way or the other, using merely the job applicants for the particular position. So for example, a plaintiff may use pass-fail rates from state or nation-wide aptitude tests to show adverse impact of a test on black males. This rests on the assumption that the larger sample includes black males who are similar enough to the plaintiffs, both in terms of ability and hypothetical desire for this job or one like it, in order to infer adverse impact on the smaller group from clear
statistical adverse impact on the larger sample. Second, the Court mistrusts applicant pools as the sole sample, due to selection bias in the decision to apply for the job. Potential employees who know about the discriminatory hiring and testing procedures will be far less likely to apply if they are minorities. While this concern is valid, the diametrically opposite approach of Dothard is problematic. Nationwide adverse impact on women does not allow for any difference between the genders in desire for prison guard jobs. If one makes the reasonable assumption that preference for such jobs is positively correlated with height, weight, and strength, regardless of gender, fewer women (particularly small women) will want to apply. Thus the statistical proof allowed in Dothard may grossly overstate the degree of adverse impact. A better approach in that case would have been to require some showing of adverse impact if these requirements were applied to other states where smaller women were currently allowed into prison guard jobs. Such a sample of women would be a large enough pool to provide inference, but would control for preference far more effectively.

From an analytical standpoint, the employer’s selection procedure begins with an applicable pool of the entire American work force. The employer publishes certain job qualifications in a geographic area or sends out word-of-mouth information about job openings. Employees also voluntarily describe their employment conditions and satisfaction (or lack thereof) to potential job candidates. All this serves to narrow the applicant pool down to a number of workers who want the job and also believe that they have a large enough chance of getting it to warrant the expenditure of their time and

---

7 Dothard v. Rawlinson, (1977). See also Griggs.
energy on the application process. Now, with an applicant pool, the firm applies further screening procedures of qualifications, tests, references, and interviews, narrowing the pool down to the number of new hires. The question is: what types of screening devices, at what points in this process, are subject to adverse impact scrutiny? In *Dothard*, the premise that the employer’s selection procedures were subject to scrutiny only after the applicant pool had been formed was rejected.

Later a court found that an employer’s requirement of previous executive experience was itself discriminatory. The very formation of the pool of applicants was held to be a violation of Title VII, and the court constructed a set of “minimum qualifications” for the job in question.

This notion that a firm need only seek some “minimum qualification” is of particular relevance to the testing issue. While it is true that certain jobs require minimum qualifications, that is not to say that an employee who has more than the minimum set of abilities or skills will not do better at the job than one who just satisfies the minimum. Especially in jobs that have some upward mobility, employers are quite reasonably interested in hiring applicants with potential for promotion within the company. For this purpose, standardized tests that yield many more than two outcomes (“qualified” and “unqualified”) are likely to be valued by employers. However, language from *Coviale* and other cases indicates that the legal system rejects this argument. An employer seems to have no general legal right to look for a “more than qualified” applicant.

---

8 Justice White made this point in his dissent in *Dothard*.

9 An interesting side point about *Dothard* is that the employer has already won, through a BFOQ defense, a claim that the “same sex only” rule for prisoners/prison guards was discriminatory. In other words, changing the height and weight requirements would only have resulted in more women (perhaps) competing for the same number of female guard jobs, not in a net increase in female positions.
The point at which courts should start scrutinizing this “narrowing down” process depends to some extent on the costs and benefits associated with this scrutiny. For example, never scrutinizing any job requirement which is advertised at the pre-applicant stage would basically nullify the anti-discrimination legislation. Employers would always find a way to classify their job requirements before receiving applications in such a way as to discourage all undesirable applicants, including minorities the employer does not want to hire. On the other hand, scrutinizing requirements of previous relevant job experience, as in Caviale, could raise the costs of hiring substantially for employers. Broad screening devices like previous executive experience are extremely cheap to implement, easy to ascertain, and useful as a first cut in eliminating candidates who are less likely to be successful. If even the “first cut” is subjected to the highest level of impact scrutiny, the number of open positions is likely to go down, harming all workers and minorities in particular. A reasonable approach could be to subject the selection procedures to an increasing degree of scrutiny as the process goes on towards actual hiring. This would recognize that the benefits of scrutinizing early steps in the process are relatively lower, while the costs to employers and workers is relatively higher. In this approach the initial advertisement or job qualification required for application would only be deemed discriminatory in more extreme cases of adverse impact and unreasonableness of the criteria, than would be the case for selection criteria applied to actual job candidates. To make this a clear-cut (bright line) rule, perhaps the courts could require a likelihood of less than 2% that the racial or gender composition of the new hires could

---

have occurred by chance, compared to the general pre-applicant population, while maintaining a 5% or 10% rule when comparing to the actual pool of applicants.

1.2.1.5 *Wards Cove* and the Civil Rights Act of 1991.

Apart from its highly controversial effects in the area of employer defenses, *Wards Cove v. Atonio* (1989) was decided at the prima facie stage and was considered good law by the Civil Rights Act of 1991. A comparison between the racial compositions of two segments of an employer's work force was not deemed to be sufficient for a prima facie case. The point was to compare the composition of the hired group with that of the applicant group, the Court said. Moreover, the plaintiffs must allege a specific selection procedure which has adverse impact, whether or not the bottom line reflects that impact. In addition, subjective employment practices, such as the alleged nepotism and separate hiring methods, do not escape adverse impact liability. Once a particular practice is shown to have adverse impact, the prima facie case is proved.

1.2.1.6 Statistical techniques and the EEOC guidelines

Although the Supreme Court itself has not dealt with complicated statistical proof cases, several courts have defined how they may be used. First, as explained above, the plaintiff may use a larger sample than the applicant pool itself to make his case. He need only show clear connection between the larger pool used and the group upon which adverse impact fell. Second, the plaintiff is to show statistical significance of the adverse impact; if limited to a smaller sample, the plaintiff is free (and encouraged) to use specialized
small-sample inference techniques\textsuperscript{11}. Third, the plaintiff needs to include in regression analysis the “major factors” affecting the proof of adverse impact, but need not include all conceivable relevant factors\textsuperscript{12}.

The Equal Employment Opportunity Commission has somewhat ambiguous interpretive rights over the anti-discrimination statutes. For the purpose of enforcement, the Commission uses the 4/5 rule, i.e. if the selection rate of any protected group falls more than 20% below that of the highest selection rate for any other group, the Commission will seriously consider prosecution of a claim against the employer. Several courts have departed from this rule, characterizing it as simply an enforcement guideline and not a recipe for liability. Some circuit courts also differ with the guidelines in measuring the group disparities, using comparisons of “pass” or “fail” rates of majority and minority workers due to a particular selection procedure\textsuperscript{13}.

1.2.2 Employer defenses

1.2.2.1 Bona Fide Occupational Qualification (BFOQ)

Under an adverse treatment theory of liability, this is the only justification available to the employer besides rebutting the intent evidence and the affirmative action defense,(which we will not deal with here). Race cannot be considered a BFOQ according to §703(e) of Title VII. As for gender, in Dothard the Court upheld a same-sex rule for prison guards and their charges on the basis of the high rate of violence towards women that was likely

\textsuperscript{11} Fudge v. City of Providence Fire Department, (1985).

\textsuperscript{12} Bazemore v. Friday, (1986).
to be exhibited by male prisoners in the state. Alabama was allowed to avoid the possible harm to female guards in male prisons, as well as the huge potential liability to itself, by forbidding female guards in male prisons. Interestingly, in *International Union v. Johnson Controls* (1991), fertile female employees succeeded in showing that a policy forbidding them from work in the lead battery area was discriminatory. This was in spite of possible harm to them and, more importantly, large potential harms and liabilities to unborn children. Although Judge Posner (at the 7th Circuit) mentions the problem of potential tort liability for Johnson Controls, the Supreme Court did not take account of cost considerations in this case, and it generally denies cost considerations for BFOQ defenses.

The two cases can possibly be reconciled because Alabama is a state employer and because not all fertile women will become pregnant. As we will discuss below, the total lack of consideration of cost to employers is troublesome, at the level of BFOQs and more so for business necessity defenses. In the face of such large potential liability, firms are likely to react in ways that the courts cannot control, such as changing production and technology, and these reactions may eliminate any beneficial effect of the restrictions.

1.2.2.2 Business necessity and job relatedness

Under *Griggs*, a firm has to prove the "business necessity" of any test or selection practice producing adverse impact. The Court quoted EEOC guidelines which indicated that an employer must display statistical evidence that this test was significantly

---

correlated with or predictive of job performance for the specific job at issue. As discussed above, the Court also gave the impression that broad testing devices were highly suspect. Later EEOC guidelines spelled out careful validation techniques for standardized tests, which confirmed this trend towards requiring job-specific tests and job-specific studies of validity.

The EEOC has split validity into criterion, construct, and content measures. The first is what we have mentioned so far, i.e. showing that the test score is a good statistical predictor for job performance taken as a whole. Construct validity requires the firm to "decompose" the characteristics of an employee that make him or her productive, and show that the test measures those specific characteristics properly. Content validity means showing that the questions or tasks of the test are a close approximation to actual tasks to be performed on the job. An entire new industry of psychometricians, test validators, and expert witnesses, has been created as a result of such guidelines. Given the prohibitively high costs of conducting such studies, an employer would prefer to abandon testing or simply hire without testing and keep employees for a trial period which amounts to a "test" of specific job ability. Alternatively, an employer could shift his attention to avoiding initial law suits by hiring according to quotas.

These guidelines point to an important observation regarding adverse impact liability and its effect on standardized testing. Nominally, subjective and objective employment practices are subjected to the same level of scrutiny. However, the types of data and
evidence that are possible to acquire with regard to standardized tests are vastly more numerous and more expensive than those for more subjective criteria. With more available “bells and whistles” in the business necessity defense for standardized tests, such defenses are far costlier than defenses of subjective practices. Therefore, rational employers will tend to bias their selection procedures towards subjective practices.

In *Wards Cove*, although the case was decided at the prima facie stage, the Court’s opinion altered the necessity defense immensely. First the Court stated that the defendant’s burden was one of production and not of proof, leaving the plaintiff to prove that a particular practice was *not* consistent with business necessity. Besides adding significantly to the plaintiff’s burden, the informational requirements of proving a negative are far more difficult to meet, especially with the inferior information available to the plaintiff. In addition, the Court said that the employer need only produce evidence that “articulated a legitimate business purpose” and expressly mentioned the uselessness of extensive validation studies for many cases in which they had been required previously. Congress was so incensed over this ruling that two years later, in the Civil Rights Act of 1991, it expressly overturned the decision and emphasized repeatedly that the Act meant to return to *Griggs* standards of business necessity.\textsuperscript{15}


\textsuperscript{15} It is interesting that two of the most influential testing cases had their greatest impact through dicta rather than holding. *Griggs* affected standardized general aptitude testing by its sweeping condemnation of “broad” procedures, and Ward’s Cove purported to almost eliminate validation studies. Perhaps this is a case of a “multitude of words” resulting in unwise statements.
I.3 Alternative employment practices

Once a plaintiff has proven the disparate impact and the defendant has rebutted with a showing of business necessity, the Court envisages a third stage in which plaintiff gets to show that an alternative test or selection procedure will have a significantly less adverse impact, while “also serving the business purpose”. The language of Albemarle Paper Co. v. Moody (1975) does not indicate cost considerations affecting the alternative business proactive showing. In other words, “serving the business purpose” cannot be interpreted as “serving the business purpose as cheaply”. The Eleventh Circuit has found that alternative business practices can be prohibitively expensive, unworkable, or unsafe.

From the language of Fitzpatrick v. City of Atlanta (1993), the degree to which cost is taken into account seems to be quite low. The Court relied far more on the facts that alternatives were probably just as uncomfortable for the firemen involved and that safety was a genuine concern. This question of the degree of cost increase that puts another practice out of the “alternative” range is still pending.

1.3 The Psychometric Perspective

1.3.1 The “best single predictor” evidence

The debate over whether there is in fact a single general intelligence quality “g” is alive and well. Respectable theories in psychology set the number of “basic” intelligence components anywhere between 1 and 716. However, for the purposes of personnel selection, psychometricians have found that the characteristics predicting job performance are highly correlated, so that an index of job aptitude maybe constructed and tested for.

Schmidt, Ones, and Hunter (1992) discuss several studies conducted using "project A" data, collected from Army testing procedures. Cognitive measures of general ability were consistently found to be the best single predictor of performance at any job requiring some complex reasoning. The authors explain that measures of general ability are nowadays constructed to be non-knowledge or skill based. They measure the test taker's ability to react to novel situations, to take responsibility and initiative in unusual circumstances, and to acquire specific skills or knowledge over time. Another study showed that once the measures of general cognitive ability were included as explanatory variables, adding other variables like education or previous experience added little to the overall explanatory power of the predicting regression\textsuperscript{17}.

In terms of the persistence over time of this predictive power, Deadrick and Madigan (1990) find that validity of general ability predictors have long term effects. They argue that this is because general ability measures learning abilities, and these have long term effects on productivity, particularly in more complex jobs. Schmidt, et. al. (1992) also consider the magnitude of benefit that such tests afford the employer. Using the same large Army sample, they split jobs into low, medium, and high complexity. They found that percentage increases in productivity based on the adoption of generalized aptitude testing were 19% for low-complexity jobs, 32% for medium complexity jobs, and 48% for high complexity jobs.

\textsuperscript{17} Ree and Earles, 1990.
In terms of bias, Schmidt et. al. (1992) cite earlier studies of the *Griggs* time period which show that aptitude tests were often biased predictors of minority performance, because the tests were more knowledge-based and required cultural and socio-economic backgrounds most minorities did not possess. However, since the days of *Griggs*, testing procedures have been upgraded considerably. For example, in the late 1980s the Labor Department Employment Service began recommending the Generalized Aptitude Test Battery for use by state employment agencies. The test resulted in adverse impact on some minority groups. However, a National Academy of Sciences report found that the test was unbiased in its prediction of minority and majority job performance and equally accurate for both groups of workers. Nevertheless, the report called for race-norming the scores due to the adverse impact. Employers in the states using the GATB objected so strenuously that use of the test battery was suspended\(^{18}\).

Add to this evidence the very low cost of standardized general aptitude testing, and you get a formidable case for allowing employers to use such general tests. The Wonderlic Personnel Test, a paper and pencil test that takes 12 minutes to complete, is said to have a correlation coefficient of .9 with "g", the elusive general ability coefficient\(^ {19}\). Separate validation studies for specific jobs and tests are vastly more expensive, causing smaller employers and firms at high risk of lawsuits to abandon testing altogether.

\(^{18}\) Hartigan and Wigdor, 1989.
\(^{19}\) Quote from Miguel Quinones, Professor of Industrial Psychology, Rice University, Houston, TX.
1.3.2 Alternative tests

Evidence on job-specific tests indicates that while they are more expensive, they measure often the very same general ability, and their use results in the same adverse impact as generalized tests. First, the evidence of specific tests’ predictive power is most often based on studies of employees already established in an occupation. It is likely, though, that an individual’s preferences determine both job choice and the set of specific skills he will devote his cognitive ability to acquiring. Those with higher ability levels are better at acquiring job specific skills and thus perform better on specific ability tests\textsuperscript{20}. Second, when used to predict performance on jobs for which they were not designed, specific ability tests are almost as good at predicting ample productivity as they are at predicting productivity in the jobs they were designed to test\textsuperscript{21}. Although employees taking tests of specific ability that were not designed for their own job scored lower than their counterparts in the correct job, the predictive power of the test was highly comparable for all jobs with the same level of complexity. The indication is that specific ability tests may often be simply a far more expensive way of measuring the same general ability and productivity. Employers use them now, not because they are technically superior, but because they are easier to defend in court.

1.4. The Economic Perspective

1.4.1 The basic informational problem

Restrictions on testing are placed on firms and workers in a world of imperfect information. Employers benefit from testing because they can better select workers and

\textsuperscript{20} Prediger (1989).
predict production. Workers benefit because they are more efficiently matched to the jobs at which they will tend to excel. Therefore, a restriction on testing is a form of informational constraint. In general, the economic perspective suspects informational constraints of being detrimental to social welfare by reducing productivity and wages and raising prices. For the purposes of this section, we will assume that we live in a world where intrinsic worker ability (or productivity) is private information held by the worker but is of value to the employer. Also, assume that jobs may be classified along a spectrum of increasing sensitivity to intrinsic ability. For lower end jobs, ability does not affect productivity much at all. For the highest level jobs, ability is crucial for productivity on the job. Firms seek to match workers with the jobs at which their level of ability is put to optimal use; that is, firms would like to allocate lower ability workers to relatively "safe" jobs, and higher ability workers to "risky" jobs.

Note that on the higher end of the spectrum, credential requirements such as degrees and specialized experience may restrict greatly the effectiveness of general ability tests as predictors of performance. Once a person has a law degree, for example, the educational screening has drastically reduced the range of general abilities he is likely to possess. Thus, giving an aptitude test to incoming law associates will not help a law firm much in predicting their performance. In other words, it is likely that the "middle" jobs are the most difficult for firms to handle; they are reasonably sensitive to ability, and yet they have no accepted credentials. Unfortunately for firms, aptitude testing procedures for such jobs are also probably the hardest to defend.

21 Ackerman and Humphreys, 1990.
1.4.2 The bases of discriminatory outcomes

In addition to an informational problem, economic analysis of the testing issue requires some pre-existing problem of discrimination. The 1964 Act’s language indicates clearly that Congress saw discrimination as an irrational practice which should be eliminated on both moral and economic grounds, and that its elimination would increase productivity and social welfare.

Economics is generally suspicious of the idea of irrationally motivated behavior. The basic paradigm of economics is that of an informed (to some extent) and rational maximizer, of utility or of profits. The objection to irrationality is perhaps mostly methodological. Assuming certain behavior is irrational means that we cannot write a rational model of choice that predicts that behavior, and so there is no use studying it at all. Among the various theories of discrimination, the closest to the “irrational” model is taste-based discrimination. Employers and perhaps customers simply “do not like” minority workers, and so the minorities get paid less and are forced into lower paying occupations. Under this theory of discrimination, the restrictions on testing may make sense because they would serve to change employer’s preferences over time. No good model of preference evolution is available, though. Also, competition in the labor and product markets would tend to eliminate this type of discrimination very quickly; this does not square with discrimination’s long life in this country.

---

Statistical discrimination\textsuperscript{23} is based on the premise that tests of worker ability are either biased or less accurate for minorities, so that the wage profile offered to them as a function of tests score is flatter. Thus, minority workers rationally decide to accumulate less human capital, because they are rewarded less for any increase in their skills or knowledge. They become truly less "able" as a result, and discrimination persists. One problem with this theory is the evidence cited above about the unbiasedness and accuracy of the current general aptitude tests. However, the idea of affecting employees incentives adversely is useful.

Discrimination based on "language" or culture is a relatively new theory\textsuperscript{24}. The idea behind it is that minority workers have a cost of communication with majority workers and managers, and since they are in the minority, they bear this cost of communication. This theory is useful because it does explain persistent discrimination without appealing to pervasive prejudice, which is unlikely, or to pervasive testing bias, which has been disproved by psychometrics.

1.4.3 Incentive effects of restrictions

Assuming that both firms and workers are rational agents seeking their best interest in profits or utility, we expect legal restriction on informational exchange to alter both agents' behavior. Firms are likely to respond in two basic ways. First, they can seek the same information about worker productivity through means that remain legal. This probably involves greater reliance on references, previous experience, and on credentials.

\textsuperscript{23} Lundberg and Startz, 1983.
However, firms are unlikely to obtain the same information after testing restrictions as they could obtain before. Certainly the same information is much more expensive to come by after testing restrictions are imposed. Therefore, firms also respond by altering their technologies in ways that mitigate the additional costs and production losses of restrictions. For example, firms may move away from “middle” jobs, using technologies that require larger numbers of low-skill workers, while using highly credentialled managers or supervisors. This alters the distribution of available jobs and could affect the income distribution among workers.

The manner in which workers’ incentives are affected depends on what theory of discrimination one subscribes to. For example, under taste-based discrimination, a minority individual cannot do anything to make majorities stop “disliking” him; so he has to live with the discriminatory outcome, or become an employer himself. Under both statistical and language based discrimination, effects on worker incentives are more substantial. Language discrimination is particularly relevant to policy makers because it claims that the cause of discrimination is at least partially controllable by the minority group itself. Since “culture” is, at least to some extent, a choice, restrictions on testing may cause minority workers to become far less willing to assimilate into the business culture of the majority, because the rewards for doing so become too uncertain due to testing restrictions.

---

24 Kevin Lang, 1986
1.5 Conclusion

If legislation against discrimination is meant to eliminate a totally irrational practice, it need not take account of cost, benefits, and incentives very carefully. We have argued that testing is neither irrational nor unproductive. Therefore, the courts and the legislature must treat the players in this game as rational agents. Legal policy makers must consider the cost and benefits involved in eliminating or restricting tests. To do this, they must understand the psychometric evidence on the productivity of testing and other selection procedures. They must also the effects on firms' selection, production, and technology decisions. Finally, they should take into account the effects on minority workers' incentives. Minorities who are encouraged to invest in learning are capable of mitigating much of the adverse effect of discrimination in their lives. If testing restrictions discourage learning and assimilation into the prevailing business culture, this anti-discrimination policy could cause more harm than help to minorities.
CHAPTER 2: Language Discrimination, Signaling, and Informational Restrictions

2.1 Introduction

Among the employment practices that have come under close scrutiny since the 1964 Civil Rights Act are the methods by which firms obtain information about worker productivity for the purposes of hiring and promotion. These methods include general aptitude testing, signals sent by workers, and other indicators of ability. In 1971, Duke Power Company required job applicants without a high school diploma to earn a minimum score on a standardized test. In *Griggs v. Duke Power Co.*\(^1\), the Supreme Court found this policy discriminatory. Blacks have lower graduation rates, so relatively more black applicants had to take the test. The fact that a high school diploma is a good indicator of important worker characteristics like determination and dependability was not considered a sufficient defense of Duke Power’s position. Such practices were thus pronounced illegal unless the employer could prove that only specific job-related skills were required or tested. As a result of this and other decisions, employment testing diminished (Johnson, 1976). In the late 1980’s, the Labor Department Employment Service recommended the General Aptitude Test Battery for use by state employment agencies (Blits and Gottfredson, 1990). Although the GATB predicted job performance equally well for blacks and whites, blacks had lower scores on average. Therefore, a National Academy of Sciences Report called for race-norming the scores (Hartigan and Wigdor, 1989).

\(^1\) 401 U.S. 424 (1971)
In its treatment of employment testing, the Supreme Court has interpreted Title VII of the Civil Rights Act of 1964 as disallowing disparate impact as well as disparate treatment. In other words, the results of a policy are on trial rather than its intent or content. Any employment practice affecting a minority group adversely is discriminatory, even if that practice treats individuals fairly.

On the other hand, in *Los Angeles Department of Water v. Manhart*\(^2\), the Court held that Title VII mandates fairness to individuals, not to groups. The issue was the use of gender as an indicator of longevity. The L.A. Department of Water based pension fund contributions from employee pay on different actuarial tables for male and female employees. The Court ruled in favor of the women employees, stating that since individual women may not share the female characteristic of outliving men, charging women more than men for the same pension constituted discrimination. The ruling did allow, however, for the Department of Water to pay its employees a fixed sum to be used for purchasing annuities from the private sector which continues to use gender-specific actuarial tables. The argument for the *Johnson Controls*\(^3\) decision is similar. Women of childbearing age could not be barred from jobs involving high exposure to lead. Note that Johnson Controls faces possible tort liability for birth defects suffered by children of female employees. This means that hiring females of childbearing age imposes an added cost on the employer. The ruling in this case implies that higher costs associated with hiring certain employees do not justify refusing to hire them.

\(^2\) 435 U.S. 702 (1978)

\(^3\) International Union, United Automobile, Aerospace and Agricultural Implement Workers of America, UAW, Et. Al. v. Johnson Controls, Inc. No. 89-1215
Differences regarding fairness to individuals or groups notwithstanding, in each of these cases the Supreme Court decision somehow limits the use of certain information. In Manhart and Johnson Controls, the use of gender as a predictor of longevity or risk of pregnancy was forbidden. And yet the fact is that only women are at risk of becoming pregnant. Also, the average woman outlives the average man, and insurance companies need information about their "average" customers. In Griggs, a test or screening device (method of gaining information) was forbidden because it adversely affected black workers. Yet, the evidence from psychometrics is that general ability tests and indicators are better at predicting worker productivity than specific, work-sample tests (Schmidt, et. al., 1992). Several questions arise concerning the results of these restrictions.

First, how do firms respond when screening is altogether forbidden? If a firm believes that a certain group of job applicants would have lower average scores if testing were allowed, then group affiliation will play a role in the hiring decision. In particular, if different jobs within a firm require different ability levels, the firm may tend to engage in more job discrimination in the absence of testing. For example, an employer may be less likely to hire minority applicants in high-ability, high-paying jobs when he is unable to test them. A greater proportion of high-ability minorities are then likely to be placed in the wrong type of job. This mismatching is only temporary provided the employer can learn an employee's potential in a high-ability position by observing that employee's performance in a low-ability position. However, workers of all ability levels may

---

4 Generally, testing or other informational restrictions are accompanied by job discrimination restrictions, such as the 4/5 rule of thumb used by the EEOC to determine whether an employer's work force is
perform alike in low-ability positions. In that case, initial mismatching is not easily corrected. Restrictions on the exchange of information may adversely affect high-ability minority workers and lower productive efficiency.

Second, how do firms and workers respond when some tests are forbidden but alternative screening methods are available? Firms wish to obtain information about job applicants, and high quality workers of all groups wish to distinguish themselves. It is therefore likely that an alternative signal or screening device will be used. Since this device is only chosen due to legal restrictions, it may be somehow inferior to the original (now illegal) test. The alternative test or screening device may be more obliquely related to gender or race, but it may also be more obliquely related to job performance. Again, as argued above, efficiency is reduced.

Third, do informational restrictions actually reduce discrimination? While it may redistribute wealth in favor of minority groups, it may reduce total wealth and may adversely affect high-ability minorities. A related question is whether or not alternative signals or screening devices impose the same costs on high-ability individuals of both minority and majority groups. Firms may have to resort to less exact and more subjective selection criteria. For example, interviews and connections may play a larger role in the selection procedure. If this is true, minority workers may be at a disadvantage when the firm's recruiters belong mainly to the majority group.

sufficiently “representative” of the qualified pool of applicants. Thus, job discrimination in response to informational restrictions could result in high legal costs for firms.
To begin exploring these issues, it is necessary to start with a basic model of labor market discrimination. We can then add worker heterogeneity in terms of ability and the notion that different workers have a comparative advantage at different jobs (worker-job matching). We must also add some restrictions on the information which can be used as a signal of high ability or gathered by firms through employment testing. It will then be possible to examine the effects of informational restrictions on efficiency and the degree of discrimination faced by minorities. It can be shown that under certain circumstances informational restrictions harm minority workers of all ability levels.

The rest of this chapter is organized as follows. Section 2.2 summarizes three models of labor market discrimination: Taste-based discrimination, statistical discrimination, and language discrimination. Section 2.3 cites some historical and empirical evidence in favor of the language model. Section 2.4 extends the language model by adding differences in worker skills, comparative advantage at different jobs, and the possibility of signaling. Section 2.5 discusses the effects of changes in legal and informational restrictions, in light of the previous section. Section 2.6 concludes and proposes some changes and additions to the model.

2.2 Models of Labor Market Discrimination

Taste-based discrimination models assume that employers or workers belonging to one group dislike individuals of some other group. Discriminators require compensation to hire or work with members of the discriminated against group. Thus, that group of workers receives lower wages (Becker, 1971; Arrow, 1972). There are two main
problems with this argument. First, tastes cannot be measured or tested. Second, the
taste for discrimination must be totally prevalent among employers or workers of the
discriminating group, or else competition will eventually drive discriminators out of the
market.

Statistical discrimination is based on the assumption that employers observe minority
workers' productivity levels with greater error than those of majority workers. Thus,
employers' estimates of productivity place more weight on the group mean for minorities,
and the wage schedule (as a function of observed productivity) is flatter for minorities. If
workers invest in human capital, minorities are compensated less for any level of
investment. Therefore, they invest less in human capital and have lower average
productivity and pay (Lundberg and Startz, 1983). The problem with this reasoning is
that it must begin with the assumption of a test or screening device which varies in
precision across groups of applicants. In the days of Griggs, standardized aptitude tests
may have been biased against minority applicants. However, in the case of the General
Aptitude Test Battery, the same National Academy of Sciences report that recommended
race-norming the test scores also confirmed that the test battery was unbiased and equally
accurate for all applicant groups (Blits and Gottfredson, 1990).

communities, blacks who speak B and whites who speak W. Workers from these two
groups are identical in every respect except language. However, there are relatively more
white capitalists, so in equilibrium some whites must hire blacks. Since learning a new
language is costly, and the total cost of hiring both types of workers must be the same in equilibrium; black wages will be depressed. In other words, the cost of communication is borne by the minority speech group. Lang goes on to consider production processes that use workers and supervisors in fixed proportions. He shows using similar reasoning that black workers are paid less than their white counterparts. Black supervisors are paid more than white supervisors, but not enough more to make it worthwhile for whites to learn B and supervise blacks. The outcome is complete segregation as well as wage discrimination. Finally, when industries differ according to worker-supervisor ratio, black workers are occupationally segregated into industries with high worker-supervisor ratios. In this case both wage and job discrimination occur (a smaller proportion of blacks can become supervisors). Our model will differ from Lang's by allowing differences in worker ability and the possibility of signaling. This enables us to examine the results of legal restrictions on information exchange in the labor market. The main new results are that informational restrictions lower efficiency, redistribute income firms to minority workers, but can harm minority workers in terms of real and relative wages.

2.3 Evidence in Favor of the Language Discrimination Model

The evidence in favor of the language model can be divided into evidence in favor of the model's motivation and that in favor of its empirical implications. First, consider some examples of separate speech communities within an economy. By the early 1800's in China, Mandarin had become the state language as well as the language of the elite. Any individual aspiring to a good position in the government (or in most private enterprises) had to be fluent in Mandarin. Since classical Mandarin is very difficult to learn, non-
native speakers were greatly disadvantaged by this situation. The language was used more as a barrier than a screening device. In British ruled India, good jobs also required knowledge of English. In that case, British entrepreneurs held much of the capital in India even though native speakers of English were a small minority in the work force. The situation in Quebec is somewhat more complicated. Within Quebec itself a large majority claims French as a native tongue; within Canada at large, native French speakers are a small minority. The geographic isolation of French speaking Canadians has both political and economic ramifications.

In America, several examples of separate speech communities exist. The most obvious example is that of immigrant communities. It appears that immigrants who stay within closely knit circles in which their own language is predominant fare worse in the labor market than their more adventurous compatriots (Chiswick and Miller, 1992). Those who never learn English well are lower-paid within their occupations and tend to settle into lower-skill occupations (Chiswick and Miller, 1995). Another instance concerns black-white language differences. A large proportion of inner city blacks speak Black English, recognized by linguists as different from Standard English (Labow, 1973). The differences between the two speech communities transcend grammar, vocabulary, and pronunciation. Other verbal as well as nonverbal differences set the two speech patterns apart. Again, inner city blacks who never become conversant in Standard English face greatly limited economic opportunities. Gender differences in communicative style have also been noted (Henley, 1977). In the U.S., women’s communicative style is generally evaluated as less effective than men’s communicative style, especially in management
positions. However, women’s style of communication is very similar to that of Japanese managers, who are considered to be quite successful with their own workers.

Several implications of the language model seem to be consistent with empirical facts. First, successful members of the minority speech community tend to become bilingual. This is certainly true among immigrants to the U.S. Second, members of the minority speech group are segregated into low-paying occupations. In the gender discrimination literature, much has been made of this very observation. Lastly, since language is generally learned in school, a minority worker will obtain more education than an otherwise identical majority worker. High-ability minorities do seem to go into occupations with clear credentialling requirements, and they generally are better qualified than majority workers for the positions they acquire.

2.4 Adding Ability Differences to the Language Discrimination Model

Assume two speech communities with Black and White as their languages. In each community, there are two types of workers, high and low ability. Ability level is private information held by the worker. The proportion of high-ability individuals is the same for both communities. As in Lang (1986), communication between the two speech groups is necessary because relatively more owners of firms are white. For simplicity, we will assume that all firm owners are white.

In each firm there are two types of jobs, worker and supervisor. The production function is
\[ Q = \min \left\{ X_s, \frac{1}{a} X_w \right\}; a > 1 \]

where \( X_s \) and \( X_w \) are effective units of supervisory and worker input. Low and high ability employees perform equally well as workers, yielding \( X_w = 1 \) per worker. In the supervisory job, high-ability employees yield 1 unit of effective labor while low-ability employees yield none. That is, \( X_s \) = number of high-ability supervisors hired by the firm.

We assume that one of the supervisor's main function is to act as "liaison" between the firm owner and the workers. Since he must relay orders from owner to workers and bring worker disputes, questions, or demands to the owner, this supervisor must be fluent in the language of the owner and that of the workers. If the supervisor cannot communicate with either the workers or the owner, then production from his "team" (a production team is one supervisor and \( a \) workers) is zero, regardless of his ability level. Thus, a black individual must know White in order to supervise in any firm, while a white individual need only know Black if he applies for a supervisory job over black workers. The supervisor also manages the production process, making decisions and solving problems that affect the production level. In addition, the supervisor may engage in monitoring of the workers to induce optimal effort expenditure by workers. The supervisor's effective contribution to output as he performs these tasks of communicating orders, organizing production, and monitoring, is highly sensitive to his ability level\(^5\).

\(^5\) As mentioned above, we make the extreme assumption that output goes to zero in a production team with a low ability supervisor. However, the qualitative results of this analysis are not affected by making a more moderate assumption that a certain percentage of output is lost due to bad supervision.
To simplify matters, we will assume that the proportion of high-ability workers in each community is $1/(1+a)$, i.e., there are just enough high-ability workers in each group to properly supervise all low-ability workers in that group. Thus, in the case of perfect information about applicant ability, all high ability individuals would be assigned supervisory positions.

Low and high ability workers also differ in terms of language learning costs. Learning the other speech community’s language costs high-ability workers $d$ and low-ability workers $d'$. Assume $d' > d$. Due to this difference in costs, learning another language may be used as a signal of high ability. If the supervisory job provides enough compensation for language learning costs in order to attract only high-ability employees, a separating equilibrium is achieved. For this to occur, the premium earned by supervisors (above worker wages) must exceed $d'$, but must not exceed $d$.

Suppose that the only screening device available to employers is a language requirement. If firms are not allowed to use language knowledge as a screen, they are forced to hire supervisors at random. When hiring one supervisor and $a$ workers at random, a firm produces 1 unit with probability $1/(1+a)$ and 0 units with probability $a/(1+a)$. On the other hand, if a firm hires a supervisor it knows to be high-ability, the final product is always 1. Therefore, the expected product gained by hiring a supervisor who has successfully signaled a high ability level is $a/(1+a)$. 
In keeping with the “business necessity” doctrine discussed earlier, we assume that an employer cannot require language knowledge for a job unless that language is necessary for the specific supervisory job. This means that high-ability whites seeking supervisory positions over white workers may not signal their ability level by learning B. An employer who requires knowledge of Black, when hiring a white supervisor to supervise white workers, runs the risk of costly lawsuits. Any white applicant for a supervisory position over whites who is turned down because he does not know Black may sue the company, claiming that the language screen is not “job-related”. Such a lawsuit would be likely to succeed, resulting in monetary losses such as back-pay, perhaps in addition to an injunction forcing the firm to hire the plaintiff into the job category in question. The only way to ensure non-liability in such a suit is to maintain a supervisory pool that represents the qualified pool of applicants. Since language knowledge is not a bona fide “qualification” unless directly related to the job, maintaining this “representative” workforce is equivalent to hiring white applicants at random into supervisory jobs over whites.

The model has only one period, and we assume that wages are fixed when the initial hiring decision is made, before actual output is realized. We envisage a large industrial production process in which total output is easily observed but output by each production team is prohibitively expensive to observe. As explained above, this model does not preclude monitoring of workers, but the quality of monitoring will depend on the ability level of the supervisor. The supervisor himself is not monitored\(^6\), and once output is

---

\(^6\) Again, due to very high costs of monitoring. In order to effectively monitor all supervisors, the firm would need to institute a third job category of “monitors”. However, once again the effectiveness of these monitors would depend on the ability of those employees, which is unknown to the firm.
known, the firm can calculate the proportion of bad supervisors it has employed, but cannot tell which supervisors are low ability.

We now examine the equilibria of this model as they relate to three parameters: \( d \), the high ability individual's cost of learning a second language, \( a \), the number of required workers per supervisor, and \( L \), the proportion of blacks in the total population. Firms seek to maximize profits and use wages as a strategy to attract the desired employees for supervisory positions. Employees' strategies are to learn the other speech group's language or not to learn it. The model outcome consists of a level of expected production, a set of wage offers, and a set of employment patterns. In an equilibrium, all observed employment patterns must yield the same expected profit\(^7\). From the employees' perspective, equilibrium implies that given the strategies chosen by all four types of individuals, each individual prefers to maintain the strategy his type has chosen. Note that in general there are four possible employment patterns: Whites supervise whites (WW), whites supervise blacks (WB), blacks supervise blacks (BB), and blacks supervise whites (BW).

**Case 1: High language learning costs, complete segregation**

\[
\frac{a}{1 + a} < d
\]  

(1)

This condition implies that the productive benefit of the signal is less than its cost for high ability individuals. In this case there is complete segregation in the work force; whites supervise whites and blacks supervise blacks. Let black and white wages for
workers be $w_b$ and $w_w$, and let $w_j^i$ denote the wage paid to a supervisor of type $i$ over workers of type $j$. Since no second language is needed to perform the supervisor’s job, $w^w = w_w$. Black supervisors must speak White to communicate with the white owner. In order to induce perfect sorting by ability among black applicants, suppose that $w_b + d \leq w_b^b < w_b + d'$. Hence, the expected product of a black firm is higher than that of a white firm, not counting the cost of language learning undertaken by the supervisors of a black firm. Normalizing the price of output to 1, the expected profits from employment patterns $WW$ and $BB$ are then:

\[
\begin{align*}
\pi_{WW} &= \frac{1}{1+a} - w^w_w - aw_w = \frac{1}{1+a} - (1+a)w_w \\
\pi_{BB} &= 1 - w^b_b - aw_b
\end{align*}
\]

The division of wage revenue between black supervisors and workers is not determined within this model. One alternative is to pay black workers their opportunity cost of working in this industry. Supposing that utility is linear and that the alternative activity outside the industry gives workers utility of 0, then the minimum wage workers can be paid while remaining in the market is 0. In that case, $w^b_b = 1$, and this equilibrium can hold as long as $d' > 1$. In this equilibrium, supervisors would be earning well above $d$ as a premium above the black worker wage. At the other extreme, we can assume that black supervisors receive no premium above the actual language learning cost, so that

\[\text{\footnote{In determining the wage levels, we will also assume that the industry is perfectly competitive, so that profit from each production team is at most zero (production teams with profit below zero will not be observed in any equilibrium).}}\]
\( w^b_w = w_b + d \). For the purposes of this paper, we will make the second assumption about the wage revenue split among blacks. In that case, setting \( \pi_{ww} = \pi_{bb} \) yields:

\[
w_b = w_w + \frac{a}{1+a} - d
\]

(1a)

So, if (1) holds, black workers earn less than white workers and black supervisory jobs pay more than white supervisory jobs, but not enough more to lure high-ability whites into learning Black. The condition for segregation to be maintained is \( w_w + d \geq w^b_b \). From (1a) we obtain

\[
w_w + d = w^b_b + \frac{d - a}{1+a}
\]

(1b)

The intuition is simple. The benefit of successful signaling is not worth the cost of learning the second language. High ability blacks learn White because it is the only way to become a supervisor. This is not true for whites, so none of them learn Black, and complete segregation results.

Note that if we had made an assumption that the equilibrium within black firms involved higher supervisor pay and lower worker pay, this would affect the condition for segregation to continue. For example, if black supervisors earn 1 and black workers earn nothing, a segregated equilibrium would only hold if \( w_w + d \geq 1^8 \).

---

\(^8\) This would imply for certain sets of parameters, that the full segregation equilibrium would break down even though \( \frac{a}{1+a} < d \). That is, although language learning is too costly a signal for the white speech group, it may be undertaken by high ability whites, due to large pay differentials between low and high ability blacks. In that sense, the equilibrium within black firms in which the black supervisor only earns a premium exactly equal to the high-ability learning cost is the most desirable. When that equilibrium
Case 2: Low language learning costs, blacks are a majority

\[ \frac{a}{1 + a} > d \]  \hspace{1cm} (2a)

\[ L > \frac{1}{2} \]  \hspace{1cm} (2b)

When (2a) holds, black workers get paid higher wages if complete segregation prevails. However, \( w_w + d < w^b_b \), so full segregation cannot be maintained. A high-ability white is now willing to incur the cost \( d \) to get a job supervising blacks. The question then arises: Can all high-ability whites obtain positions supervising blacks? The answer depends directly on \( L \), the proportion of blacks in the work force. If \( L > 1/2 \), then all high-ability whites can supervise blacks and thus signal their ability successfully. In this equilibrium, employment patterns WB, BW, and BB are observed. All high-ability whites supervise blacks, and high-ability blacks supervise the rest of the black workers as well as all white workers. This is the case of a completely separating equilibrium; every supervisory job is filled by a high-ability individual. In this equilibrium, wages for workers and supervisors are equal across groups. To see this, consider employment patterns BB and BW. The expected profits of these production teams are:

\[ \pi_{BB} = 1 - w^b_b - aw^b_b \]

\[ \pi_{BW} = 1 - w^b_w - aw_w \]

Black supervisors in both types of firms successfully signal their high ability. In equilibrium, they must earn the same wages, or else black supervisors of blacks will undercut black supervisors of whites, or vice versa. But if \( w^b_b = w^b_w \), then \( w_b = w_w \).
Notice that whites learning Black is an unproductive activity from a technical point of view. Yet whites learning B does improve the matching of workers to jobs, increasing overall expected product.

**Case 3: Low language learning costs, blacks are a minority**

\[
\frac{a}{1+a} > d \quad (3a)
\]

\[
L < \frac{1}{2} \quad (3b)
\]

\[
d \leq \frac{aL}{1-2L+a-aL} \quad (3c)
\]

Given (3a), as argued above, high-ability whites would like to get jobs supervising blacks. Unfortunately, due to (3b) there are not enough black workers to go around. Supervisory jobs over black workers are now rationed, so that white supervisors of blacks are chosen randomly from all whites who have learned B\(^9\). In order for all high-ability whites to learn B when only some of them can supervise blacks in the equilibrium, the pay differential between white supervisors of whites and white supervisors of blacks must exceed \(d\). Only \(L/(1-L)\) proportion high-ability whites can get jobs supervising blacks.

To make the expected benefit of learning B positive for high-ability whites requires

\[
w_b^w = w_w + \frac{d(1-L)}{L} \quad (3d)
\]

\(^9\) A firms only interviews for language knowledge among white applicants until it has filled all supervisory jobs over black workers. It is never worthwhile to interview an additional applicant because any information gained about his ability cannot be used effectively. If a white applicant who knows black is placed in a white supervisory job, he firm cannot legally defend the decision to hire him for that job as opposed to another white applicant who does not speak Black.
If (3a) and (3b) are true and if all high-ability whites find it worthwhile to learn B in order to get a chance to supervise blacks, we would expect the equilibrium to exhibit employment patterns WW, WB, and BB. All high-ability whites learn B and as many as possible supervise blacks, all high-ability black workers supervise whites, and the leftover white workers are supervised by whites.

Note that black high ability individuals can successfully signal their ability while supervising black or white workers. Therefore, this equilibrium will only hold if \( w^b_w \geq w^w_w \). To find equilibrium wages, compare employment patterns WW and WB.

Since \( \frac{L}{1 - L} \) proportion of high ability whites get jobs supervising blacks, the remaining proportion of high ability individuals in the pool available to WW production teams is \( \frac{1 - 2L}{1 - 2L + a - aL} \). Expected profits from employment patterns WW, WB, and BB therefore are:

\[
\pi_{WW} = \frac{1 - 2L}{1 - 2L + a - aL} - w^w_w - aw^w_w
\]

\[
\pi_{WB} = 1 - w^w_b - aw^w_b
\]

\[
\pi_{BB} = 1 - w^b_w - aw^b_w
\]

As in case 1 above, white supervisors of whites cannot signal and are paid \( w^w_w = w^w_w \).

Using (3d) and setting all expected profits equal to zero yields values for wages:

\[10\] To see this, subtract \( \left( \frac{L}{1 - L} \right) \left( \frac{1}{1 + a} \right) \left( 1 - L \right) \) from the original numbers of high ability whites and all whites, which are \( \frac{1 - L}{1 + a} \) and \( 1 - L \) respectively, and find the new proportion.
\[
    w_w = \frac{1-2L}{(1+a)(1-2L+a-aL)}
\]
\[
    w_b^w = \frac{1-2L}{(1+a)(1-2L+a-aL)} + \frac{d(1-L)}{L}
\]
\[
    w_b = \frac{1}{a} \left\{ 1 - \frac{1-2L}{(1+a)(1-2L+a-aL)} - \frac{d(1-L)}{L} \right\}
\]
\[
    w_w^b = 1 - \frac{a(1-2L)}{(1+a)(1-2L+a-aL)}
\]

In order for high-ability blacks not to undercut white supervisors of blacks, it must be the case that \( w_w^b \geq w_w^w \). This condition is captured by (3c)\(^{11}\). The figure below shows the combinations of parameters that result in the above cases, given a certain value of \( a \).

We plot \( d \), the high-ability cost of learning a new language, against \( L \), the proportion of blacks in the population at large. In Region I, (1) holds, so the equilibrium is that of
perfect segregation with black workers earning less than white workers and black supervisors earning less than white supervisors, once language learning costs are accounted for. In Region II, (2a) and (2b) are true; the completely separating equilibrium of case 2 occurs. In this area, wages for supervisors and workers are equal across groups\textsuperscript{12}. In Region III, (3a-c) are satisfied and black workers and supervisors earn at least as much as their white counterparts. Finally, (3c) is violated in Region IV, so there is no equilibrium in which all high-ability whites make the same learning decision. As the number of workers per supervisor grows, regions III and IV change as shown below. With more workers per supervisor, the set of \((L,d)\) combinations that result in no equilibrium with a pure learning strategy becomes larger with respect to the area of pure strategy equilibria.

\begin{center}
\textbf{Regions III/IV for different worker/supervisor ratios}
\end{center}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{regions.png}
\end{figure}

\textsuperscript{11} For pure strategy as well as for mixed strategy equilibria, this condition is an upper bound on \(d\). For values of \(d\) that are too high, it is impossible to compensate high ability whites enough to induce all of them to learn Black, while still paying black supervisors of whites a higher wage.

\textsuperscript{12} Parameters that would generate such outcomes are not likely. It is generally true that the majority speech group, defined in this model as the group whose members own all the firms, is also a numerical majority in the labor force.
Case 4: A fraction $k$ of high-ability whites learn Black

\[
\frac{a}{1 + a} > d \quad (3a)
\]

\[
L < \frac{1}{2} \quad (3b)
\]

\[
d \leq \left( \frac{aL}{1 - 2L + a - aL} \right)^\frac{1}{k} \quad (4)
\]

This leads to the consideration of mixed strategy equilibria, in which a certain proportion $k$ of high-ability whites learn B and the rest do not. Equilibrium in this case requires

\[
k \geq \frac{L}{1 - L} \quad ; \text{i.e. more than enough high ability whites must learn Black in order to supervise all low ability blacks. To see this, consider the hypothetical equilibrium if this condition is violated. If not enough high ability whites learn Black in order to supervise all low ability blacks, some black workers must be supervised by blacks. All four employment patterns would then be present. In this equilibrium, it must be true that } w^b_b = w^w_w, \text{ since employment patterns } BB \text{ and } WB \text{ must yield the same expected profit.}
\]

Also, $w^b_b \geq w^w_w + d$, or else no high ability white would have learned Black to begin with.

Since there are low ability black workers that are not being supervised by whites, an additional high ability white who learns Black can obtain a supervisory job over blacks, by undercutting slightly a black supervisor over black workers. Since the premium earned by a black supervisor over the white wage is at least $d^{13}$, additional high ability whites will learn Black, until all possible black supervisory positions are filled by high ability whites.
With enough high ability whites learning Black in order to supervise all black low ability workers, the three employment patterns BW, WB, and WW are observed in mixed strategy equilibria. The same proportion of high ability whites is missing from the WW employment pool, so expected profits and wages for such production teams are the same as in case 3. Given $w_w$ from WW, $w^*_w$ is determined as in case 3, and black supervisory wages are unchanged. With a lower proportion of high ability whites learning Black, the probability of receiving a supervisory job over blacks, for each white who learns, is higher than in the pure strategy equilibrium. Therefore, the premium earned by white supervisors of blacks is not as high. In order for a proportion $k$ of high ability whites to be properly compensated for learning, we need $w^*_w = w_w + \frac{dk(1-L)}{L}$. This implies

$$w^*_w = \frac{1-2L}{(1+a)(1-2L+a-aL)} + \frac{dk(1-L)}{L}$$

$$w^*_b = \frac{1}{a} \left\{ 1 - \frac{1-2L}{(1+a)(1-2L+a-aL)} \right\} \frac{dk(1-L)}{L}$$

$w^*_w \geq w^*_b$ is now captured by (4). The figure below compares, for $a=2$, the range of parameters for which a pure strategy equilibrium is feasible with the range for which an equilibrium is feasible $k=.5$. As $k$ falls, the set of feasible parameters increases in size.

---

13 Recall that within the WW employment pattern, supervisors and workers must be paid alike, since no signaling is available, and supervisors must be selected at random from the pool of applicants.
Below $d = \frac{a}{1+a}$, it is always possible to find a mixed strategy equilibrium. In the following figure we show the sets of feasible parameters for pure strategy equilibria and for mixed strategy equilibria in which exactly enough high ability whites learn Black to supervise all low ability blacks. Note that the condition on $d$ for such mixed equilibria to be feasible is within region I, i.e. whenever complete segregation breaks down as an equilibrium, it is always possible to find a mixed strategy cross-supervision equilibrium as in region IV\textsuperscript{14}.

\textsuperscript{14} To see this algebraically, replace $k = \frac{L}{1-L}$ into (4). The resulting condition is $d \leq \frac{a(1-L)}{(1+a)(1-L) - L}$, which is always true with $L>0$ and outside of region I.
Note that in each of these equilibria, black workers and supervisors are paid at least as much as white workers and supervisors, respectively. However, in order for more high ability whites to learn Black, wages of bilingual white supervisors must rise, and black worker wages must fall to keep the WB employment pattern profitable. Thus, a policy to increase integration by inducing more Black learning by high ability whites may result in lower wages for black workers, without increasing the pay of black supervisors. Such a policy would harm the black speech group as a whole. In addition, it would lower overall welfare, because expected compensation for a high ability white does not change (bilingual white supervisory wages rise, but the probability of receiving that wage falls), while actual compensation for low ability blacks falls. The more high ability whites learn Black above the proportion needed to supervise all black workers, the lower overall welfare. This is because, given the legal restriction on using Black as a signal in a WW production team, additional learning by whites is not productive.
2.5 Effects of Legal Restrictions

The above equilibria indicate that under some plausible circumstances minority workers can benefit from the freedom firms have to gather information about general ability through signaling. We now consider the effects of two changes in legal restrictions on the outcomes of this model.

To begin with, suppose we relax the restriction on whites earning a premium for knowing B when they are supervising white workers. In this scenario signaling will be freely undertaken by high-ability members of both speech groups as long as \( a/(1 + a) > d \). In region I, language learning is too expensive, so it is not used as a signal by the white community even without legal restrictions. In region II, the signal is used regardless of the restriction, because all high ability whites can find jobs supervising black workers. As mentioned above, this situation, in which the minority speech group is a numerical majority, is unlikely to occur. Reality in the U.S. probably resembles regions III and IV: relatively low language learning costs between "speech communities", and minority workers are a numerical minority in the work force. Relaxing the legal restriction in these regions will result in all high ability whites learning Black and successfully signaling their ability. This improves sorting in the labor market, equalizes wages of workers and supervisors across groups, and raises overall welfare. As in case 1 above, we need an additional assumption about how wage revenue is split between workers and supervisors to pin down the wages for all types of labor. If supervisors from both groups are compensated for language learning costs but do not earn any other surplus, then removing legal restrictions will benefit white and black low ability individuals, at the expense of
high ability blacks. This implies that the most likely result of policies of legal restrictions is lower overall welfare and greater income disparity within the minority group. In other words, the legal restriction on signaling may increase high ability minority income at the expense of low ability income in both minority and majority communities.

Another change in restrictions could be to forbid the use of language as a screening device. Since all blacks must learn a language to become supervisors, whereas whites need not learn a language, blacks may seem to be at a disadvantage from a legal perspective. As a result, firms are no longer allowed to impose any language requirements on potential supervisors of either speech group. In this model, supervisors must therefore be hired at random, and firms must then provide them with language training (without being able to find out the language learning cost of an individual supervisor). Complete segregation will result in this situation, and black supervisor wages will be lower than in case 1 above. This is because black supervisors must now pay, through lower wages, for the average language learning cost in the black population, not that of high-ability blacks. The addition of this restriction lowers productive efficiency and profits, and worsens the situation of the black speech group. High-ability blacks are harmed because they are not allowed to signal. More surprisingly, low-ability blacks can be harmed because their productivity depends on supervisor ability. With less efficient sorting, supervisors are less likely to be high ability, and worker wages suffer.
2.6 Conclusion

We find that adding differences in worker ability changes the results of the original language discrimination model significantly. Language becomes a productive signal; if this signal is not too costly, both speech groups use it in order to improve employee/job sorting and increase productive efficiency. In the absence of any legal restriction on its use as a signal, this tends to equalize wage rates between speech groups. When the use of this signal is restricted to "job related to the specific job in question" instances, productive efficiency and overall welfare unambiguously decrease, and high ability black wages may increase at the expense of low ability wages for blacks and whites. A more severe restriction, such as prohibiting the use of language knowledge as a screen for any job, lowers productivity even more and lowers all wages while maintaining wage equality between groups. Thus, high and low ability employees from both communities are harmed by such a policy.

One shortcoming of this analysis is that firms are not agents with a meaningful choice. They have no way of responding to restrictions other than abiding by them. This problem is addressed in the next chapter by giving firms a choice of technology. Also, this model assumes that signaling through language knowledge is the only alternative to random supervisory hiring. We will now include a standardized test in addition to the language signal. This provides an alternative screening device, and allows us to study the case of more than one legal restriction. We consider the "job related" restriction on language knowledge as an employment screen as well as similar restrictions on standardized testing that affect the accuracy of the test.
CHAPTER 3: Standardized Testing Restrictions and Employment Discrimination

3.1 Introduction

The employment testing controversy began in 1971 with *Griggs v. Duke Power Co.*\(^1\). Without finding Duke Power Co. guilty of any intent to discriminate, the Supreme Court ruled that an employment selection procedure that results in adverse impact on any minority group is discriminatory unless the employer can prove that the selection procedure is a “direct business necessity”. In particular, the Court argued that general aptitude testing procedures operated as “built-in headwinds for minority groups” and were poor measures of potential job productivity. The Court considered general ability tests to be bad predictors of job performance as well as biased against minority workers. Over the past two decades, various types and amounts of evidence have been required to establish a prima facie case of discrimination based solely on the adverse impact of a standardized testing procedure. There have also been several “legal regimes” governing the evidence a firm must present in defense of its testing procedure once a prima facie case has been established. Nevertheless, under all legal regimes, general aptitude testing as an employment selection procedure has been greatly restricted and is much more costly when it is used (Johnson, 1986).

Meanwhile, in the psychological literature on mental ability testing, evidence in favor of general aptitude tests as good predictors of job performance has been mounting. Psychometricians have found that measures of general cognitive ability are usually the

\(^1\) 401 U.S. 424 (1971).
most valid predictors of productivity on jobs that contain some mental tasks, and that their validity as predictors can be generalized across firms, jobs, and occupations. Such measures are especially important as predictors when workers who do not have any experience are initially sorted into jobs and occupations. The cost of administering standardized tests that measure general cognitive ability is very low (disregarding costs of validation studies and litigation costs imposed by legal restrictions).

In other words, standardized aptitude testing appears to be the cheapest and the most effective way to gather information that is crucial to the employment selection process. It is therefore important to examine the effects of testing restrictions on the behavior of firms as well as the choices made by workers. Informational restrictions may affect the choices made by agents in the labor market. As a result, they may affect output, firm profits, and wages earned by individuals in both the majority and minority groups.

Lang (1986) proposes a model of discrimination based on language differences. “Language” means all verbal and non-verbal modes of communication. Differences in language include all cultural differences affecting individual communication. In Lang’s model there are two speech groups forced to interact in the labor market because relatively more owners of firms belong to the majority group. Communication between owner and employees is necessary for production, and communication across speech groups is costly. This cost is borne by members of the minority speech group in the form of lower wages. In addition, when technologies differ according to worker-supervisor ratios, minorities are segregated into firms with high worker supervisor ratios. Sadka
(1996) extends this model by including high- and low-ability workers within each speech group. Language learning costs are lower for high-ability individuals, so language can be used as a signal of ability level. It is shown that under certain plausible circumstances, this signaling process results in higher wages for all members of the minority speech group. Eliminating the use of language as a screening device lowers output and profits and can also lower the relative wages of both high and low ability minorities.

This paper undertakes two main extensions of Sadka (1996). First, we focus on restrictions placed on standardized testing as a selection procedure. This is accomplished by including a standardized test that imperfectly measures general cognitive ability. Firms now gather information about potential employees by administering a standardized test and by conducting interviews that reveal language knowledge. Legal restrictions alter the accuracy of the standardized test which firms are allowed to use. Second, we allow firms a choice of technology. Technologies differ in terms of worker-supervisor ratios and in terms of how sensitive output is to the ability level of the supervisor. We find that as standardized test become less accurate, production unambiguously falls. More importantly, high ability applicants are more likely to benefit from testing restrictions than are low ability workers. In particular, among minority language groups the higher ability employees tend to benefit from increased test error, at the expense of low ability majority group workers.

Section 3.2 provides background and motivation for the approach taken in this paper by appealing to literature in the areas of law and psychometrics. Section 3.3 considers a
model with discrete testing outcomes. Section 3.4 explores an extension of the model to continuous testing outcomes, and discusses preliminary results. Section 3.5 concludes and offers extensions.

3.2 Background and Motivation

*Griggs v. Duke Power Co.* was the first case involving standardized employment testing to come before the Supreme Court. Before this case, proving employment discrimination involved providing evidence of disparate treatment, i.e. intent to discriminate. It was in *Griggs* that the Court first held that disparate impact without intent to discriminate could constitute valid proof of employment discrimination. Prior to the effective date of Title VII, Duke Power Company hired blacks only into the lowest paying of its five departments. The company also required a high school diploma for all potential employees in the other four departments. Once Title VII came into effect, Duke Power Company began to require satisfactory scores on two professionally developed aptitude tests in order to promote or hire employees into the four upper level departments. Both selection procedures adversely affected the hiring and promotion of black workers. The district and circuit courts found no discriminatory intent in the adoption of the education and test requirements, and ruled that Duke Power Company had not violated Title VII. The Supreme Court concurred on lack of discriminatory intent, but reversed the decision, ruling that selection procedures which have adverse impact on minorities are illegal unless the employer can prove "direct business necessity."
Prior to *Griggs*, the Equal Employment Opportunity Commission had published its first "Guidelines on Employee Selection Procedures."\(^2\) This publication provided guidelines for proper validation of employment tests as job related. Although these guidelines were stringent interpretations of Title VII, they did not affect employment testing greatly because the courts did not treat them as imperatives. This changed in *Albemarle Paper Co. v. Moody*\(^3\). The company had conducted a professional validation study of its testing procedures. However, the Supreme Court found Albemarle Paper Company guilty of employment discrimination because its validation methods did not meet the EEOC’s guidelines.

The legal regime governing employment testing has varied greatly in recent years. In *Wards Cove Packing Co. v. Atonio*\(^4\) the Court held that, in response to a prima facie case of disparate impact, the employer need only articulate “a legitimate business purpose.” The Court later interpreted this to mean that technical validation studies of testing procedures were not always necessary. The controversy caused by this reversal resulted in the Civil Rights Bill of 1991. This legislation returns to *Griggs* standards. In addition, it requires a local validation study for each job in each company. Obviously this imposes large costs on firms using employment tests.

One example of the legal controversy over employment testing is the General Aptitude Test Battery. In the late 1980s, the Labor Department Employment Service

\(^2\) 29 C.F.R. Sec. 1607.1-.14 (1975)  
\(^3\) 422 U.S. 405 (1975).  
recommended the GATB to state employment agencies (Sadka, 1996). On average, black workers scored lower on this test. This brought about a National Research Council Committee investigation of the GATB. The committee found that general cognitive ability measures were equally valid predictors of job performance regardless of race (Hartigan and Wigdor, 1989). Nevertheless, the Labor Department decided to race-norm the scores. Employers’ strong objections resulted in a complete suspension of the GATB for use by state employment agencies.

While the legal climate has been generally hostile to employment testing, research in psychometrics continues to gather evidence in favor of general aptitude tests as selection procedures. In the late 1980s, the Army conducted the “largest and most expensive selection research project in history” (Schmidt et al., 1992). This study, called Project A, found measures of general cognitive ability to be the best single predictor of job performance in occupations requiring any complex reasoning. Using the same data, Ree and Earles (1990) found that adding other variables to cognitive ability increased explanatory power very slightly.

These findings are important because the use of general cognitive ability as a screening device seems to have substantial and long term effects on output. Deadrick and Madigan (1990) find the validity of general cognitive ability as a predictor to be stable or increasing over job tenure. Hunter et. al. (1990) use several available studies to estimate the percentage increases in output due to the use of general aptitude testing. Categorizing jobs according to complexity level, they estimate the percentage increases as follows:
19% for low-complexity jobs, 32% for medium-complexity jobs, and 48% for high-complexity jobs.

Psychometrics has also gathered evidence concerning alternatives to standardized general ability testing, such as tests of specific abilities. These tests are suspect on two counts. First, the evidence in favor of their predictive power is built mostly on data concerning workers already sorted into jobs and professions. However, it is likely that an individual’s preferences determine both job choice and the specific skills he will devote his general cognitive ability to developing. Those with higher ability have higher levels of specific skills and perform better on the job (Prediger, 1989). Second, specific ability tests seem to be merely a more expensive way of measuring general cognitive ability. Ackerman and Humphreys (1990) consider the predictive power of specific ability tests for performance in jobs which do not require those abilities. They find that such tests are equally good predictors for all jobs with the same general level of complexity. Another important alternative selection procedure is interviews. Dipboye et al. (19??) find that the predictive power of interviews varies greatly with the interviewer. This suggests that more able interviewers are much better at predicting job performance. It is not unlikely that more able interviewers are a more costly resource in terms of output.

3.3 Model of Discrete Ability and Testing Outcomes

We focus on a simple model structure of discrete ability levels and test outcomes. We assume that ability is intrinsic and binary. A proportion p of the population is high ability. In this paper we study the case of two test outcomes; the results and intuition are
similar for a larger number of discrete outcomes. Employers receive a test score for each applicant, and for the case of two speech groups, they also interview those applicants who must know a non-native language to perform their prospective jobs. In each of the cases studied below, we consider the effects of increasing test error through legal or informational restrictions.

3.3.1 Technologies

Two technologies are available to the industry in perfectly divisible proportions. They are defined as

\[
Q_1 = \min[a_1 X_s, a_2 X_w^1] \\
Q_2 = \min[d_1 X_m, d_2 X_w^2]
\]

where \(X_s\) is effective supervisory units, \(X_m\) is effective managerial units, \(X_w^1\) is effective worker units assigned to \(Q_1\), and \(X_w^2\) is effective worker units assigned to \(Q_2\).

We also assume:

\[
\frac{a_1}{a_2} > \frac{d_1}{d_2}
\]

\(X_w^1\) = number of workers hired in \(Q_1\).  

\(X_w^2\) = number of workers hired in \(Q_2\).  

\(X_s\) = number of high ability supervisors employed in \(Q_1\).

---

5 We do find that total production increases invariably as the test outcomes become more numerous, even though ability itself is binary. This is contrary the common wisdom in the legal system. 

6 This assumption is based on the notion that a technology allowing less responsibility to each manager (so that not all output is destroyed by a low ability manager) would require more managers per worker.
\[ X_m = \text{number of high ability managers} + (1/2) \times \text{number of low ability managers} \]

employed in \( Q_2 \).

This means that \( Q_1 \) uses more workers per supervisor but is more sensitive to the ability level of the supervisor. Imagine that each supervisor or manager and his workers are a production team; while each bad supervisor destroys all of the output from his team, a bad manager will only destroy half of the output. In other words, \( Q_1 \) represents a technology in which the supervisor has more responsibility and more freedom to innovate. Hence the supervisor may potentially jeopardize more production than the manager.

The reasoning behind these assumptions concerning technology is twofold: Firms face restrictions on: the type of information they can gather, the types of tests they can give to prospective and current employees, and the mix of workers they can hire in any job category. One of the few ways in which a firm can respond to such restrictions is via its choice of technology. A simple way of allowing technology choice to respond to restrictions is to introduce free movement between two fixed proportions technologies. In the real world there are certainly adjustments within each type of technology; but we would expect these adjustments to be small compared to changes in the type of technology. We isolate these technological “shifts” by eliminating any adjustments within technologies. As shown below, each point in the set of convex combinations of \( Q_1 \) and \( Q_2 \) will represent a specific mix of the “dangerous” and “safe” technologies.
Another reason for including two technologies is that the "middle jobs" are the scene of the greatest controversy. Consider a spectrum of jobs ordered according to increasing sensitivity of productivity to the general ability of employees. As you move to the right, jobs become complex and employees are faced with more novel, non-routine problems. Psychometricians point out that jobs on the high end of this spectrum tend to require agreed upon credentials. Once these requirements have been met by prospective employees, general ability testing is not very useful because the applicants are all well above average. On the low end of the spectrum, productivity is only slightly dependent on general ability, so general ability tests are also of limited use. It is the jobs in the middle that pose the real problem. Without rigid credentialling, applicants can vary greatly in terms of general ability, and this affects productivity significantly. Therefore, firms stand to gain from general ability testing. However, establishing the job-relatedness of general ability test items is difficult for such jobs, since they require only limited background knowledge or training. In other words, the jobs for which testing is most useful are precisely the jobs for which testing is greatly hampered by legal restrictions.
One potential problem with the fixed proportions technologies is that firms may choose to be at one end of the segment connecting $Q_1$ and $Q_2$, i.e. at a corner solution technologically speaking. This can be remedied by choosing benchmark production parameters so that, in perfect information, the firm chooses to be strictly in the interior of the set of convex combinations of the two technologies.

### 3.3.2 Standardized Testing

Suppose the standardized test yields only two outcomes, $\alpha$ and $\beta$. Test accuracy can be illustrated as follows:

![Diagram of test outcomes](image)

- $\alpha_L$ and $\beta_H$ capture the error of the test. From $\alpha_L$ and $\alpha_H$ we obtain the conditional probabilities $P(H/\alpha)$ and $P(L/\alpha)$, and from $\beta_L$ and $\beta_H$ the conditional probabilities $P(H/\beta)$ and $P(L/\beta)$. $P(L/\alpha)$ represents the likelihood of false positives, while $P(H/\beta)$ is the likelihood of false negatives. If we assume that high ability individuals are a minority and that test errors are symmetric, meaning that $\alpha_L=\beta_H$, then false positives will be more plentiful than false negatives. The comparison between the perfect information case and a case with test error can be illustrated by

---

$^7$ The assumption of symmetric error is made for simplicity. However, it is possible that test writers are more concerned with avoiding Type I or Type II error, and would adjust $\alpha_L$ and $\beta_H$ accordingly.
This discrepancy between false positives and false negatives increases as the symmetric test errors become larger, and increases faster when high ability individuals are a smaller minority of the work force. This means that the expected product from using a high-scoring employee as either a supervisor or a manager is reduced. However, as production in the supervisory technology is more sensitive to ability level, the expected product of a supervisor will be more adversely affected by test error than that of a manager with the same score.

3.3.3 Benchmark Case: Perfect Information, One Speech Group

The choice of technology in this case can be obtained as a solution to the cost minimization problem for the industry\(^8\). The planner solves the following problem\(^9\):

\[
\text{Min } w_\alpha p + w_\beta (1 - p)
\]

---

\(^8\) We normalize the number of employees and the output to 1. Since we do not address the size of the firm, this analysis actually considers the effects of testing restrictions on an entire industry or on a wide range of jobs. One assumption that may not fit well with the industry analogy is that the supply of workers is fixed, and every worker must be placed in a job.

\(^9\) For this perfect information case, we keep the notation of \(\alpha\) and \(\beta\) scores; there is no test error, so \(\alpha\) is synonymous with high ability and \(\beta\) with low ability.
s.t. \[ \left( \frac{1}{a_1} + \frac{1}{a_2} \right) w_{\alpha} \geq 1 \]

\[ \frac{1}{d_1} w_{\alpha} + \frac{1}{d_2} w_{\beta} \geq 1 \]

\[ \frac{1}{d_1} w_{\alpha} + \frac{1}{d_2} w_{\beta} \geq 1 \]

\[ \left( \frac{1}{d_1} + \frac{1}{d_2} \right) \left( \frac{1}{2} \right) w_{\alpha} \geq 1 \]

The technologies represented in these constraints are zero profit conditions for H/H in \( Q_1 \), H/L in \( Q_1 \), H/L in \( Q_2 \), and L/L in \( Q_2 \). In order that the perfect information choice of the planner be a combination of H/L in \( Q_1 \) and H/L in \( Q_2 \), the following conditions must hold:

\[ \frac{a_1 a_2}{a_1 + a_2} < \frac{d_1 d_2}{d_1 + d_2} \]

\[ \frac{a_1 a_2}{a_1 + 2a_2} > \frac{d_1 d_2}{d_1 + 2d_2} \]

\[ \frac{d_1}{d_2} < \frac{N_H}{N_L} < \frac{a_1}{a_2} \]

Given these parametric conditions, imagine a perfect test that yields a score of \( \alpha \) if the applicant is high ability, and a score of \( \beta \) if the applicant is low ability. The perfect information problem can be drawn as:
The minimum industry labor cost is achieved at point A. The planner chooses to assign each high ability person to a supervisory or managerial job, and chooses the combination of technologies needed to accomplish that goal. In the simulations below, we choose parameters that satisfy the above constraints, so that both technologies are used (although in different proportions) as the test error changes.

3.3.4 Imperfect Information, One Speech Group

We now add uncertainty about true employee type. Each worker knows his own true ability, but the only way to convey this to the industry is to take an imperfect test of ability, as described above. We assume that the test is costless, so that all workers regardless of ability level will take it. The industry problem becomes

\[
\text{Min} \quad w_\alpha P(\alpha) + w_\beta P(\beta)
\]

s.t. \[ \left( \frac{1}{a_1} + \frac{1}{a_2} \right) \frac{w_\alpha}{P(H \mid \alpha)} \geq 1 \]
The expected zero profit conditions for this problem show that the employer can now only calculate a conditional probability that an applicant is high or low ability, based on their test score. In supervisory technologies, no output is produced when the applicant is truly low ability; hence only the conditional probability that he is high ability matters. Under managerial technologies, production is reduced but not eliminated when the manager is low ability.

We use benchmark parameters of $p=1/3$, $a_1=5.2$, $a_2=2$, $d_1=4$, $d_2=2.4$. We solve this problem and simulate it over test error levels. Starting with a test error of .03 and increasing test error yields the following solutions.
We are particularly interested in the equilibrium wages for high and low scoring individuals, the level of production, and the output shares of the two technologies. These are shown below.

As expected, the production level falls as information is restricted. More mistakes in the sorting of employees into high and low ability jobs means that a larger proportion of managers and supervisors are not equipped to perform their jobs. This damages not only their own production, but also that of all members of their “production team”. The graph of output shares shows that the industry will adjust to larger testing error by moving out
of the supervisory technology and into the managerial technology. Note that this adjustment does not eliminate the production losses, although it mitigates the damage.

![Equilibrium Wages, p=1/3](image)

The effects in terms of wage levels and relative wages are less predictable. While the original "interior" set of technologies is used, high scoring individuals experience wage increases as a result of greater uncertainty, while low scoring applicants suffer. This result is intriguing because employees with $\alpha$ scores earn higher wages even though the likelihood of an $\alpha$ score being high ability is lower. The reason is that the industry, to reduce its production loss from lower quality sorting of employees, shifts into a technique of production that requires a larger proportion of high scoring applicants. Hence $\alpha$ scores become a more scarce resource, and their compensation increases.

In fact, once uncertainty is added to the test, the industry would prefer to move completely into the safer managerial technology, and $Q_2$ is preferred more strongly the
higher the degree of uncertainty. However, not enough $\alpha$ scores are created by the test for full movement into $Q_2$. Although more $\alpha$'s occur with higher test error and the managerial technology is used more intensively, as long as there are not enough $\alpha$'s to shift completely into $Q_2$, manager compensation continues to increase. For each level of test error, the wages for a smaller test error will not clear the market for $\alpha$'s and $\beta$'s. In order to reduce industry demand for $\alpha$'s and clear the market, a higher wage for high-scoring individuals is necessary. In other words, the effect of increasing test error is to make the distribution of income more dispersed. Thus, as a policy to help low-scoring workers, restricting the test (increasing its error) is likely to fail.

Once the industry has shifted completely out of $Q_1$, wages for $\alpha$ scores begin to fall. In this range of uncertainty, the industry uses $\alpha/\beta$ in $Q_2$ and $\beta/\beta$ in $Q_2$. As test error increases, the latter technology becomes more profitable (as there are more false negatives) while the former becomes less profitable. Therefore, high-score wages must fall relative to low-score wages in order to clear the market. Finally, when test error increases even more, there are so many $\alpha$ scoring applicants that the industry uses $\alpha/\alpha$ in $Q_1$, placing high scoring applicants in worker jobs. Since a $\beta$ scoring applicant can do the worker job just as well as an $\alpha$ scoring applicant, at this point wages equalize between the two groups. All surplus associated with high ability level or high test score is thus eliminated with enough uncertainty.
The industry's adjustment to lower test accuracy can be illustrated by the following diagram.

As uncertainty increases, several changes occur in the original technologies diagram. The isocost line becomes flatter as more applicants score high. The zero profit lines for each technology also shift. For all technologies except $\beta/\beta$ in $Q_2$, this shift is inward, indicating that wages must decrease to maintain zero profits. However, this result is reversed for $\beta/\beta$ in $Q_2$, because this technology depends totally on low scoring applicants. As false negatives become more plentiful, the quality of the low scoring pool rises, so that expected output from that technology increases. Also, the shifts in zero profit lines for $Q_1$ technologies are more pronounced, as their expected output is more strongly affected by uncertainty. As shown by the movement from point A to point B, while $\alpha/\beta$ in $Q_1$ and $\alpha/\beta$ in $Q_2$ are used, high-scoring wages increase both in absolute terms and relatively to low scoring wages.
For proportions of high ability individuals greater than 1/3, the test error almost immediately causes a switch into the $\alpha/\beta$ in $Q_2$ and $\alpha/\alpha$ in $Q_1$ combination, making wages for both scores equal. This is because, at a relatively low level of uncertainty, $\alpha$ scores become so common that the industry must place them in worker jobs in order to use all individuals in production. At this point, the marginal $\alpha$ scoring applicant is performing a job in which productivity is independent of ability level or test score; consequently this worker and all other $\alpha$ scoring individuals are not compensated for the higher likelihood that they are high ability.

For $p=1/4$, we find that the range of test error for which the $\alpha/\beta$ in $Q_1$ and $\alpha/\beta$ in $Q_2$ combination is used is much smaller. Hence the effect of increasing $\alpha$-score wages is short lived. This can be seen in the diagram for technology changes in the $p=1/4$ case, shown below.
The technology adjustment is from point A through point B and eventually to point C. Except for a very small range between .05 and .07, \(\alpha\)-score wages fall throughout the range of uncertainty considered (see figure below).

![Equilibrium Wages, p=1/4](image)

Note that original perfect information wages for high ability individuals increase as they become a smaller proportion of the population. Additional scarcity drives up the rents they earn from production. This means that they have more to lose from the introduction of testing restrictions that increase test error. However, since high ability persons are so scarce to begin with, the degree of uncertainty required to eliminate their premium (by making their wages equal to those of low scoring applicants) is considerably higher (test error of .25 in the p=1/4 case as compared with test error of .13 in the p=1/3 case).

Therefore, when they are a scarcer resource in the underlying ability distribution, high ability individuals sustain more consistent losses over the test error spectrum, but they retain a wage premium for scoring \(\alpha\) up to a greater degree of uncertainty.
3.3.5 Imperfect Information, Two Speech Groups

We now introduce an additional speech group to the population. The second speech group has the same underlying ability distribution as the original group considered above. We assume that regardless of ability level or test score, minorities cannot supervise or manage without knowing the majority speech pattern, and majorities cannot supervise or manage minorities without knowing their speech pattern.

We call this the minority speech group because it is a numerical minority and because all owners of firms (or production teams) belong to the majority group. Hence, in a world of homogenous ability levels, the minority language group bears the cost of language learning, in the form of lower wages (Lang, 1986). Once heterogeneity in ability is added to this model, however, minority wages can rise because minorities use language knowledge as a signal of ability level. Since majority job applicants cannot use language knowledge as a signal unless they are interviewing for a job supervising minority workers, minority wages for both supervisors and workers can earn a premium over their majority group counterparts.

In this paper we add not only heterogeneity in ability, but also standardized testing as an initial screen of ability level. The employer has a choice of two selection procedures, in whatever order he chooses: a costless but imperfect test of underlying ability, and a costly interview that perfectly reveals language knowledge. Given a high enough cost of
interviewing and a test which is accurate enough, the industry will only interview those who receive high scores in the standardized test\textsuperscript{10}.

Note that legal rules do not allow an employer to give hiring preference or higher pay to an applicant that holds a skill or ability not directly useful for the job he is hired for. In other words, bilingual majority supervisors or managers who oversee majority workers must not be paid a higher wage than monolingual majority applicants with the same test score. Thus it is never in the interest of the industry to interview majority supervisors or managers of majority workers.

In view of these assumptions, the hiring sequence is as follows.

1. Before entry into the labor market, individuals of both speech groups decide whether or not to learn the other group's language. Learning is less costly for high ability individuals, and within an ability category, learning costs are equal across speech groups. At this point, individuals know their own true ability level, the costs of language learning, the accuracy level of the test, and the proportions of true high and low ability individuals.

2. All job applicants (the entire population) take the standardized test, which is unbiased against minority applicants and equally accurate for both groups.

3. Job applicants for bilingual supervisory or managerial positions (including all minority applicants, and majority applicants for supervising or managing minority workers) receive an interview.

\textsuperscript{10} This condition was checked by allowing the firm to choose a technology in which it interviews low-scoring applicants, as well as a technology in which it assigns them to jobs without interviewing them. For
4. Those who pass the interview receive a managerial or supervisory job and are paid the wage associated with their test score. Applicants who do not pass the interview are labeled low ability and receive the low-score wage.

Therefore, for a minority population, the employer would have to use both test and language interview to solve the following problem:

\[
\begin{align*}
\text{Min} & \quad w_\alpha P(\alpha) + w_\beta P(\beta) \\
\text{s.t.} & \quad \left(\frac{1}{a_1} + \frac{1}{a_2}\right)\left(w_w + d_i \frac{1}{P(H | \alpha)} \right) \geq 1 \\
& \quad \left(\frac{1}{a_1} w_\alpha + d_i \frac{1}{P(H | \alpha)} \right) + \frac{1}{a_2} w_\beta \geq 1 \\
& \quad \left(\frac{1}{d_1} \left(\frac{1}{d_2} \frac{1}{P(H | \alpha)} \right) + \frac{1}{d_2} w_\beta \right) \geq 1 \\
& \quad \left(\frac{1}{d_1} + \frac{1}{d_2}\right) \frac{w_\beta}{P(H | \beta) + \frac{1}{2} P(L | \beta)} \geq 1
\end{align*}
\]

where \(d_i\) is the cost of the interview. The problem statement differs from the one speech group case in that the firm incurs an additional cost of interviewing all \(\alpha\) applicants. For each true high ability \(\alpha\) score it finds, the firm must interview \(\frac{1}{P(H | \alpha)}\) applicants with

\[
\text{values of interview cost above .05 and for test error below .33, the industry does not find it worthwhile to interview \(\beta\) scoring applicants.}\
\]
score \( \alpha \). However, once an applicant receives a high score and passes the interview, he is known to be high ability\(^{11}\).

### 3.3.6 Enforced Segregation

The two speech group imperfect information case can result in a number of equilibria, which depend on the parameter settings for technology, proportion of minorities, costs of language learning, and interview costs. We do not examine all possible equilibria in this paper, but focus on two relevant types of equilibrium. The first of these is enforced segregation. This case is shown below for the \( p = \frac{1}{3} \) case\(^{12}\).

<table>
<thead>
<tr>
<th>Discrete Case: Two Speech Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha ) wage</td>
</tr>
<tr>
<td>( \beta ) wage</td>
</tr>
<tr>
<td>Product</td>
</tr>
<tr>
<td>Prod. using q1 tech.</td>
</tr>
<tr>
<td>Prod. using q2 tech.</td>
</tr>
<tr>
<td>Tech. 1: prod. share</td>
</tr>
<tr>
<td>Tech. 2: prod. share</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Minority Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha ) wage</td>
</tr>
<tr>
<td>( \beta ) wage</td>
</tr>
<tr>
<td>Product</td>
</tr>
<tr>
<td>Prod. using q1 tech.</td>
</tr>
<tr>
<td>Prod. using q2 tech.</td>
</tr>
<tr>
<td>Tech. 1: prod. share</td>
</tr>
<tr>
<td>Tech. 2: prod. share</td>
</tr>
</tbody>
</table>

---

\(^{11}\) In this problem statement, an additional technology that interviews \( \beta \) scoring applicants for language knowledge is allowed. However, the firm never chooses to interview low scoring applicants unless the degree of uncertainty is extremely large.

\(^{12}\) A more complete version of this table is provided in Appendix 2.
The first relevant result is that, for low levels of test error, there is no incentive for high ability majority group members to learn the minority speech pattern. Thus, if segregation is not enforced, only minority applicants would like to be bilingual at low levels of uncertainty. For this range, minority supervisors and workers receive lower wages than their majority counterparts. This is puzzling because language knowledge acts as a perfect signal among minorities, so that supervisors and managers among minorities are known to be high ability. Meanwhile, those holding the same positions in majority production teams are paid more although their ability level remains uncertain. The reason behind this is again relative scarcity of the resource involved, and the resulting technology choice by the industry. Among the minority group, as uncertainty increases, a smaller proportion of applicants pass both the test and the interview. Thus, the supervisory technology is used heavily, and high ability individuals become a less scarce resource. This depresses the wages of high ability supervisors and managers. Minority workers earn a higher wage than majority workers, because they are a scarcer resource among minorities. The relationship between the two high score wage levels can be seen below.
The heavy use of technology Q1 and the production levels are also shown here. Note that production is higher among minorities, because the firm is able to obtain perfect information about the ability level of high scoring minority applicants.

3.3.7 Integration is allowed.

Given that applicants can choose to apply for jobs supervising or managing workers from the other speech group, integration will occur over large ranges of uncertainty considered here. At a low level of test error, it is always in the interest of the high ability minority employee to apply for a job supervising or managing majority workers. His ability level is known to be high because he speaks the majority language; hence the industry benefits from paying him the same wage as a majority supervisor, and admitting him to a job supervising majority workers. If minorities are a very small proportion of the population, all high ability minorities who score $\alpha$ can obtain supervisory or managerial jobs among majority workers without altering the combination of technologies used for majority workers. For larger proportions of minorities, no equilibrium in pure strategies (Learn other language vs. Don’t learn other language) is available for low test error levels.
This situation prevails until a large enough test error causes high scoring majorities to lose all their wage premium above their workers. In the example above, this happens at a test error of 11%. For lower values of \( p \) (lower high ability proportion), a larger level of test error is required to trigger this change. Since high scoring individuals cannot earn less than the low scoring, wages are equalized among all majority employees. At this point, the segregation equilibrium may hold if the cost of learning another language is relatively high for high ability individuals of both speech groups. For example, in the case shown above a language learning cost of .4 would maintain a segregated equilibrium between test errors of .11 and .17. High ability minority applicants will find it worthwhile to learn the majority speech pattern in order to obtain a non-worker job; however, high ability majority applicants will not find it worthwhile to learn the minority speech pattern in order to obtain a job supervising or managing minorities.

For this range of uncertainty, if language learning costs are low enough, high ability majority applicants will learn the minority language and apply for jobs among minority workers. A limited number of high scoring minorities can be interviewed and assigned to minority workers. This follows from the assumption that the minority group is a numerical minority as well as language-disadvantaged. The resulting equilibrium involves the maximum possible amount of cross-supervision. All high ability and \( \alpha \)-scoring minorities manage majority workers, while all high ability \( \alpha \)-scoring majority employees learn the minority language and apply for jobs among minorities. All those high ability majority applicants who can be placed in managerial jobs among minorities obtain them; the rest must supervise majority workers and cannot be interviewed or
rewarded for language knowledge. Notice that the managerial technology is the only one used for bilingual production teams, because this maximizes the proportion of high ability majority applicants who can effectively signal their true ability. A larger degree of signaling increases productivity, so the shift towards $Q_2$ increases overall production.

<table>
<thead>
<tr>
<th>Employment Patterns</th>
<th>Segregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology Usage</td>
<td></td>
</tr>
<tr>
<td>$a/b$ $q_1$ for all high ability min.</td>
<td>$a/b$ $q_1$ and $a/b$ $q_2$ for majority</td>
</tr>
<tr>
<td>CL</td>
<td>0.02</td>
</tr>
<tr>
<td>Majority $\alpha$ wage</td>
<td>1.9868803</td>
</tr>
<tr>
<td>Majority $\beta$ wage</td>
<td>1.15506434</td>
</tr>
<tr>
<td>Minority $\alpha$ wage</td>
<td>2.1993273</td>
</tr>
<tr>
<td>Minority $\beta$ wage</td>
<td>1.15506434</td>
</tr>
<tr>
<td>Majority W/B man. pay</td>
<td>N/A</td>
</tr>
<tr>
<td>Total product</td>
<td>1.4674889</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Employment Patterns</th>
<th>Integration (increases as error rises)</th>
<th>Segregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology Usage</td>
<td>$a/b$ $q_1$ for BW $a/b$ $q_2$ for WW $a/b$ $q_1$ and $a/b$ $q_2$ for WW</td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Majority $\alpha$ wage</td>
<td>2.20144462</td>
<td>2.26662424</td>
</tr>
<tr>
<td>Majority $\beta$ wage</td>
<td>0.96789727</td>
<td>0.90773147</td>
</tr>
<tr>
<td>Minority $\alpha$ wage</td>
<td>2.68345769</td>
<td>2.63868619</td>
</tr>
<tr>
<td>Minority $\beta$ wage</td>
<td>0.79780524</td>
<td>0.7619006</td>
</tr>
<tr>
<td>Majority W/B man. pay</td>
<td>2.67017924</td>
<td>2.7301856</td>
</tr>
<tr>
<td>Total product</td>
<td>1.4357823</td>
<td>1.428349</td>
</tr>
</tbody>
</table>

The table above shows the solution to the two speech group case for levels of test error from 2 to 9% (using parameters: $a=.34$ for both speech groups, $a_1=5.5$, $a_2=2$, $d_1=4$, $d_2=2.4$, $L =$ proportion of black applicants = .2, $d =$ cost of learning for high ability = .4). As test error increases, production falls as before. Integration only occurs when language learning costs are low enough compared with the benefits of learning for high ability majority applicants. The high ability strategy in for these levels of test error is assumed to be learning the minority language with .4 probability. If we were to consider a strategy of learning the minority language with probability 1, integration would only occur for test
error of .7. That is, as test error increases, this allows equilibria with higher proportions of high ability majorities learning the minority language. Thus, increased error tends to increase integration (until we reach a corner at test error of .8). As integration increases, wages of minority supervisors increase, while low ability minorities are made worse off. This indicates that a policy of restricting tests in order to increase integration in the work force could have regressive effects on the minority income distribution.

3.4 Continuous testing outcomes, One Speech Group

For this version of the standardized testing model, we continue to assume a binary ability level that is relevant to job productivity. However, each individual is endowed with another ability, best seen as a ‘test-taking skill’, which is correlated imperfectly with true underlying productive ability.

We assume test-taking ability $\alpha$ such that $\alpha \sim N(0, \sigma_\alpha^2)$. For $\alpha > \alpha^*$, individuals are high-ability as described previously, while for $\alpha < \alpha^*$ they are low-ability.

The general ability test returns scores

$$\gamma = \alpha + \varepsilon$$

with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$

$$\Rightarrow \gamma \sim N(0, \sigma_\alpha^2 + \sigma_\varepsilon^2)$$

Constructing a continuous model of test scores is motivated by two main factors. First, general cognitive ability levels are considered by psychometricians to be normally
distributed, as are other natural characteristics, such as height. For this reason, tests of
general ability or aptitude are constructed to yield normally distributed scores, and to be
unbiased predictors of ability levels. This would imply a test similar to the one described
here. Second, the continuous version of this model is useful for studying the benefit to
employers of using a test that does not measure productive ability directly, but measures
another ability correlated with productivity. The legal system in the United States has
claimed repeatedly that if an employer need only know whether an applicant is qualified
or not qualified, that employer may not use a finer measure of aptitude to select
employees from a group of applicants that are all “qualified” for the job. In other words,
if the employer cannot show that person A is qualified for the job and person B is
unqualified, he may not hire A rather than B on the grounds that A is “more qualified”.

In this model, the employer is using a test that returns scores \( \gamma \), which are unbiased
predictors of a true underlying aptitude \( \alpha \). However, this underlying aptitude is not
exactly job productivity. If an employer could ascertain the exact \( \alpha \) of an applicant, he
would not care whether that person scored a \( \gamma \) of 20, 50, or 100. However, not knowing
the true \( \alpha \), the employer does care about how high a person’s score is. The higher a
person’s score, the more likely it is that this person’s true \( \alpha \) exceeds \( \alpha^* \). Thus, an
expected profit maximizing firm will always prefer an employee with a higher score.
There is no cutoff score \( \gamma \) for which a firm will be willing to say “everyone above this
score is qualified, and we will not prefer persons with higher scores within this group”.
In other words, a finer measure of this “correlated ability” is productive for an employer.
In this context, we can study the results of testing restrictions that inhibit the employer's access to measures of abilities correlated with productivity.

Note that once imperfect information is introduced, the optimal cutoff scores that admit employees into worker, managerial, and supervisory jobs necessarily change. To see this, perform the following exercise. Starting with the benchmark case, assume that all high-ability individuals are employed as supervisors or managers. Now make true ability level private knowledge held by employees and introduce the test. Suppose that firms choose to hire as supervisors and managers all those whose expected ability level exceeds \( \alpha^* \). The test score required will then be

\[
\gamma^* = \frac{\alpha^* (\sigma_a^2 + \sigma_e^2)}{\sigma_a^2}.
\]

However, fewer applicants will score above \( \gamma^* \) than those who have true ability above \( \alpha^* \). In other words, the firms will have to alter the mix of technologies which they use. This depends on the high-low ability cutoff being above the mean of the ability distribution (i.e. on high ability individuals being a minority). The degree to which the technological mix is altered is positively related to the variance of the test error. One simple way of modeling testing restrictions is raising this error variance.

In this paper we consider only the one speech group continuous test score case. Here prospective employees have no meaningful choice. They simply take the standardized test and receive wage offers for different job categories. They choose their best offer.
Firms choose score cutoffs for managers and supervisors in order to maximize expected profits, taking wages as given. We assume workers' wages to be \( w_w \), and we allow managers' and supervisors' wages to vary with test scores. They are paid \( w_m(\gamma) \) and \( w_s(\gamma) \) respectively. Those who score below \( \gamma_m \) will receive managerial and supervisory wage offers which are below the workers' wage rate, and so they choose to be workers. The best wage offer received by those with \( \gamma_m < \gamma < \gamma_s \) is for a managerial job; the rest become supervisors. This wage schedule is depicted below.

\[\text{wages} \]
\[\begin{array}{c}
W_m(\gamma) \\
W_s(\gamma) \\
W_w \\
\end{array}\]
\[\begin{array}{c}
\gamma_m \\
\gamma_s \\
\text{test scores} \\
\end{array}\]

Firms also choose the number of employees they will hire at each test score level. Since firms behave competitively, this choice implies zero profit conditions which allow us to find \( w_m(\gamma) \) and \( w_s(\gamma) \) as functions of the workers' wage rate. Imposing an equilibrium condition in the market for workers pins down their wage rate.
The industry now solves the following profit maximization problem:

$$
\max_{\gamma_s, \gamma_m, n(\gamma)} \int_{\gamma_s}^{\gamma_m} (a_1 F(\gamma - \alpha^*) - w(\gamma))n(\gamma)d\gamma + \int_{\gamma_m}^{\gamma_s} \left( d_1 \frac{1}{2} (1 + F(\gamma - \alpha^*)) - w(\gamma) \right)n(\gamma)d\gamma
$$

$$
- w_w \left[ \frac{a_1}{a_2} \int_{\gamma_s}^{\gamma_m} n(\gamma)d\gamma + \frac{d_1}{d_2} \int_{\gamma_m}^{\gamma_s} n(\gamma)d\gamma \right]
$$

The first two terms indicate expected product (which is expected revenue since price is normalized to 1), minus payments to supervisors and managers. Notice that the expected product of any given supervisory production team is \( a_1 F(\gamma - \alpha^*) \). With probability \( F(\gamma - \alpha^*) \), the supervisor is high ability, and the firm gets \( a_1 \) in product from his team. With probability \( 1 - F(\gamma - \alpha^*) \), the supervisor is low ability, and the firm gets nothing from his team. For managers the calculation is similar, except that only half of the expected production is lost if the manager is low ability.

The resulting first order conditions are:

\( w(\gamma_0) = d_1 \frac{1}{2} (1 + F(\gamma_0 - \alpha^*)) - w_w \frac{d_1}{d_2} \) for \( \gamma_m \leq \gamma_0 \leq \gamma_s \)

\( w(\gamma_0) = a_1 F(\gamma_0 - \alpha^*) - w_w \frac{a_1}{d_2} \) for \( \gamma_0 \geq \gamma_s \)

\( F(\gamma_m - \alpha^*) = 2 w_w \frac{d_1 + d_2}{d_1 d_2} - 1 \)
\[ F(\gamma_s - \alpha^*) = \frac{\omega_w \left( \frac{a_1}{a_2} - \frac{d_1}{d_2} \right) + \frac{1}{2} d_1}{a_1 - \frac{1}{2} d_1} \]

The first two conditions are zero profit conditions for supervisory and managerial production teams. After workers are paid in each team, the manager or supervisor’s compensation is determined as the remainder of the expected product of that team. Therefore, managers and supervisors with higher test scores are paid more. The last two conditions determine the cutoff scores \( \gamma_s \) and \( \gamma_m \) as functions of the worker wage rate.

The intuition behind the conditions is as follows. For supervisory units, the firm will set the cutoff score such that the expected profit of hiring a person with that score as a manager or a supervisor is exactly equal. We expect that as test error variance increases, firms will set the supervisory cutoff higher and higher, because the supervisory technology becomes more “risky”. For the managerial technology, the cutoff is set at the exact point where the residual compensation that the firm would be willing to pay to the manager equals the worker wage. A manager with a lower score would have to accept a wage below the worker’s wage, which is impossible, since an applicant with any test score can obtain at least the worker wage.

Due to the normality assumption regarding \( \alpha \) and the test error distribution, this model cannot be solved analytically\(^{13}\). We simulate this model for different sets of technology parameters and for different proportions of high ability individuals. In each simulation we consider the effects of increasing the variance of the test error, \( \varepsilon \), by imposing testing

\(^{13}\) Cumulative distribution functions for normal distributions do not exist analytically.
restrictions. Our benchmark parameters are $\alpha^* = 0.4$, $a_1 = 5$, $a_2 = 2$, $d_1 = 4$, and $d_2 = 2.4$. We consider an ability distribution that is standard normal, and alter the variance of the error.

The benchmark case shown below indicates that supervisor wage at the supervisory cutoff increases due to uncertainty, as in the discrete case.

<table>
<thead>
<tr>
<th>Continuous Case: $a_1 = 5$, $\alpha^* = 0.4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test error ($\sigma_e$)</td>
</tr>
<tr>
<td>Test score st. dev. ($\sigma_T$)</td>
</tr>
<tr>
<td>Equilibrium worker wage $W_w$</td>
</tr>
<tr>
<td>Supervisor wage at $\gamma_e$</td>
</tr>
<tr>
<td>Optimal manager cutoff $\gamma_m$</td>
</tr>
<tr>
<td>Optimal supervisor cutoff $\gamma_s$</td>
</tr>
<tr>
<td>Proportion low ability</td>
</tr>
<tr>
<td>Proportion of workers</td>
</tr>
<tr>
<td>Proportion of managers</td>
</tr>
<tr>
<td>Proportion of supervisors</td>
</tr>
<tr>
<td>Prob $\gamma_s$-score is high ability</td>
</tr>
<tr>
<td>Prob $\gamma_m$-score is high ability</td>
</tr>
<tr>
<td>Production / supervisory tech.</td>
</tr>
<tr>
<td>Production / managerial tech.</td>
</tr>
<tr>
<td>Total output</td>
</tr>
<tr>
<td>Managerial share of output</td>
</tr>
<tr>
<td>Supervisory share of output</td>
</tr>
<tr>
<td>Relative wages $W_w / W_d$</td>
</tr>
</tbody>
</table>

![Equilibrium worker wage $W_w$](image1)

![Supervisor wage at cutoff](image2)

Note that this effect ends at a large enough degree of uncertainty. This is equivalent to the discrete case in which all the possible shifting into technology Q2 has taken place,
and high scoring wages are depressed accordingly. In spite of the shifting in technology, which raises the cutoff scores for manager and supervisors, the probability of a high ability manager or supervisor at their respective cutoffs continues to decline as uncertainty increases. Therefore output falls over all possible ranges of test error. These results are illustrated below.

Appendix 3 shows other cases of this continuous model. These cases show that if high ability individuals are intrinsically more scarce in the population, they will tend to benefit from testing restrictions over a larger range of test error. This tends to increase the importance of the "regressive" effects of testing restrictions.

3.5 Conclusions

Several important questions regarding testing restrictions remain to be answered by this research. Nevertheless, the models and results discussed here indicate some important conclusions. First, testing restrictions are always likely to result in lower output, causing the entire "size of the pie" for the restricted industry to diminish. This is because testing restrictions reduce matching efficiency and lowers the expected productivity of all supervisors and managers. As such, any claim that testing restrictions result in no
efficiency losses is suspect. Second, the tradeoff involved in implementing testing restriction does not result, in a one speech group world, in any benefits to the low ability or low-skilled worker. Quite the opposite. Restrictions tend to be a regressive policy, because firms will adjust to restrictions by moving to safer technologies that are less worker intensive. Third, in the case of two language groups, increasing testing restrictions can increase the degree of integration in the work force. However, this does not better the lot of low ability minorities. Such policies benefit high ability minority employees at the expense of low ability minority workers. The result is a regressive effect on the income distribution in both speech groups.
### Appendix 1: p=1/4

#### One Speech Group: p=1/4

<table>
<thead>
<tr>
<th>Technology Usage</th>
<th>$\alpha$ in q1</th>
<th>$\beta$ in q1</th>
<th>$\alpha$ in q2 and $\beta$ in q2</th>
<th>$\alpha$ in q1 and $\beta$ in q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>QL</td>
<td>0.03</td>
<td>0.05</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td>Prop. score $\alpha$</td>
<td>0.233</td>
<td>0.275</td>
<td>0.266</td>
<td>0.236</td>
</tr>
<tr>
<td>Prop. score $\beta$</td>
<td>0.735</td>
<td>0.725</td>
<td>0.715</td>
<td>0.706</td>
</tr>
<tr>
<td>$\alpha$ wage</td>
<td>2.949473</td>
<td>2.153846</td>
<td>2.315766</td>
<td>2.252479</td>
</tr>
<tr>
<td>$\beta$ wage</td>
<td>0.757663</td>
<td>0.944066</td>
<td>0.789474</td>
<td>0.773996</td>
</tr>
<tr>
<td>Product</td>
<td>1.398485</td>
<td>1.276748</td>
<td>1.224474</td>
<td>1.210106</td>
</tr>
<tr>
<td>Prod. using q1 tech.</td>
<td>1.30375</td>
<td>1.169231</td>
<td>0.994008</td>
<td>0</td>
</tr>
<tr>
<td>Prod. using q2 tech.</td>
<td>0.004736</td>
<td>0.107517</td>
<td>0.220466</td>
<td>1.210106</td>
</tr>
<tr>
<td>Tech. 1: prod. share</td>
<td>0.994632</td>
<td>0.915788</td>
<td>0.811784</td>
<td>0</td>
</tr>
<tr>
<td>Tech. 2: prod. share</td>
<td>0.003538</td>
<td>0.084212</td>
<td>0.188216</td>
<td>1</td>
</tr>
</tbody>
</table>

#### Production at p=1/4

![Production Curve](image)

#### Output Share by Technology

![Output Share Chart](image)
## Appendix 2: Two speech group discrete case

### Discrete Case, Majority Group

<table>
<thead>
<tr>
<th>Technology Usage</th>
<th>$\alpha$ in q1</th>
<th>$\alpha$ in q2</th>
<th>$\beta$ in q1</th>
<th>$\beta$ in q2</th>
<th>$\alpha / \beta$ in q1 and $\alpha / \beta$ in q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Ql}$</td>
<td>0.03</td>
<td>0.05</td>
<td>0.07</td>
<td>0.09</td>
<td>0.11</td>
</tr>
<tr>
<td>Prop. score $\alpha$</td>
<td>0.3496</td>
<td>0.3560</td>
<td>0.3624</td>
<td>0.3688</td>
<td>0.3752</td>
</tr>
<tr>
<td>Prop. score $\beta$</td>
<td>0.6504</td>
<td>0.6440</td>
<td>0.6376</td>
<td>0.6312</td>
<td>0.6248</td>
</tr>
<tr>
<td>$\alpha$ wage</td>
<td>2.0675</td>
<td>2.2014</td>
<td>2.3307</td>
<td>2.3673</td>
<td>1.3548</td>
</tr>
<tr>
<td>$\beta$ wage</td>
<td>1.0915</td>
<td>0.9879</td>
<td>0.8468</td>
<td>0.7864</td>
<td>1.3548</td>
</tr>
<tr>
<td>$\text{Product}$</td>
<td>1.4327</td>
<td>1.4070</td>
<td>1.3857</td>
<td>1.3694</td>
<td>1.3548</td>
</tr>
<tr>
<td>$\text{Prop. sups. in q1}$</td>
<td>0.0726</td>
<td>0.0543</td>
<td>0.0380</td>
<td>0.0368</td>
<td>0.0001</td>
</tr>
<tr>
<td>$\text{Prop. mass. in q2}$</td>
<td>0.0770</td>
<td>0.0317</td>
<td>0.0264</td>
<td>0.0062</td>
<td>0.0749</td>
</tr>
<tr>
<td>$\text{Prop. workers in q1}$</td>
<td>0.1887</td>
<td>0.1411</td>
<td>0.0936</td>
<td>0.0614</td>
<td>0.0002</td>
</tr>
<tr>
<td>$\text{Prod. using q1 tech.}$</td>
<td>0.3500</td>
<td>0.2561</td>
<td>0.1650</td>
<td>1.5654</td>
<td>0.0004</td>
</tr>
<tr>
<td>$\text{Prod. using q2 tech.}$</td>
<td>1.0767</td>
<td>1.1569</td>
<td>1.2224</td>
<td>0.0130</td>
<td>1.3544</td>
</tr>
<tr>
<td>$\text{Tech. 1: prod. share}$</td>
<td>0.5485</td>
<td>0.3820</td>
<td>0.1179</td>
<td>0.0000</td>
<td>0.0003</td>
</tr>
<tr>
<td>$\text{Tech. 2: prod. share}$</td>
<td>0.7515</td>
<td>0.8160</td>
<td>0.8821</td>
<td>1.0000</td>
<td>0.0997</td>
</tr>
</tbody>
</table>

### Discrete Case, Minority Group

<table>
<thead>
<tr>
<th>Technology Usage</th>
<th>$\alpha$ in q1</th>
<th>$\alpha$ in q2</th>
<th>$\beta$ in q1</th>
<th>$\beta$ in q2</th>
<th>$\alpha / \beta$ in q1 and $\alpha / \beta$ in q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Ql}$</td>
<td>0.03</td>
<td>0.05</td>
<td>0.07</td>
<td>0.09</td>
<td>0.11</td>
</tr>
<tr>
<td>Prop. score $\alpha$</td>
<td>0.3496</td>
<td>0.3560</td>
<td>0.3624</td>
<td>0.3688</td>
<td>0.3752</td>
</tr>
<tr>
<td>Prop. score $\beta$</td>
<td>0.6504</td>
<td>0.6440</td>
<td>0.6376</td>
<td>0.6312</td>
<td>0.6248</td>
</tr>
<tr>
<td>$\alpha$ wage</td>
<td>1.7511</td>
<td>1.7468</td>
<td>1.7425</td>
<td>1.7379</td>
<td>1.7323</td>
</tr>
<tr>
<td>$\beta$ wage</td>
<td>1.2857</td>
<td>1.2857</td>
<td>1.2857</td>
<td>1.2857</td>
<td>1.2857</td>
</tr>
<tr>
<td>$\text{Product}$</td>
<td>1.4742</td>
<td>1.4703</td>
<td>1.4664</td>
<td>1.4625</td>
<td>1.4588</td>
</tr>
<tr>
<td>$\text{Prop. sups. in q1}$</td>
<td>0.1201</td>
<td>0.1146</td>
<td>0.1080</td>
<td>0.1017</td>
<td>0.0969</td>
</tr>
<tr>
<td>$\text{Prop. mass. in q2}$</td>
<td>0.3007</td>
<td>0.1744</td>
<td>0.1480</td>
<td>0.1250</td>
<td>0.0957</td>
</tr>
<tr>
<td>$\text{Prop. workers in q1}$</td>
<td>0.3358</td>
<td>0.3663</td>
<td>0.4366</td>
<td>0.4873</td>
<td>0.5379</td>
</tr>
<tr>
<td>$\text{Prod. using q1 tech.}$</td>
<td>0.6715</td>
<td>0.7226</td>
<td>0.8726</td>
<td>0.9746</td>
<td>1.0751</td>
</tr>
<tr>
<td>$\text{Prod. using q2 tech.}$</td>
<td>0.8028</td>
<td>0.8927</td>
<td>0.9528</td>
<td>0.9879</td>
<td>0.9830</td>
</tr>
<tr>
<td>$\text{Tech. 1: prod. share}$</td>
<td>0.4555</td>
<td>0.5255</td>
<td>0.5957</td>
<td>0.6664</td>
<td>0.7374</td>
</tr>
<tr>
<td>$\text{Tech. 2: prod. share}$</td>
<td>0.5445</td>
<td>0.4745</td>
<td>0.4043</td>
<td>0.3336</td>
<td>0.2626</td>
</tr>
<tr>
<td>Cost of interview</td>
<td>0.0350</td>
<td>0.0356</td>
<td>0.0362</td>
<td>0.0369</td>
<td>0.0375</td>
</tr>
</tbody>
</table>
Appendix 3: Continuous case, one speech group

a1=6.5 case: supervisors require more workers.

<table>
<thead>
<tr>
<th></th>
<th>Test error (σ²)</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test score st. dev. (σγ)</td>
<td>1.1662</td>
<td>1.2006</td>
<td>1.4142</td>
<td>1.4866</td>
<td>1.5620</td>
<td>1.6401</td>
<td>1.7205</td>
<td>1.8028</td>
<td>1.8868</td>
<td></td>
</tr>
<tr>
<td>Equilibrium worker wage W_w</td>
<td>1.2646</td>
<td>1.2406</td>
<td>1.2309</td>
<td>1.2286</td>
<td>1.2273</td>
<td>1.2267</td>
<td>1.2266</td>
<td>1.2268</td>
<td>1.2273</td>
<td></td>
</tr>
<tr>
<td>Optimal manager cutoff γ_m</td>
<td>0.6909</td>
<td>0.7172</td>
<td>0.7616</td>
<td>0.7899</td>
<td>0.8188</td>
<td>0.8510</td>
<td>0.8851</td>
<td>0.9208</td>
<td>0.9579</td>
<td></td>
</tr>
<tr>
<td>Optimal supervisor cutoff γ_s</td>
<td>1.1340</td>
<td>1.2438</td>
<td>1.5628</td>
<td>1.6747</td>
<td>1.7879</td>
<td>1.9023</td>
<td>2.0176</td>
<td>2.1336</td>
<td>2.2504</td>
<td></td>
</tr>
<tr>
<td>Proportion low ability</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td>0.6554</td>
<td></td>
</tr>
<tr>
<td>Proportion of workers</td>
<td>0.7232</td>
<td>0.7123</td>
<td>0.7049</td>
<td>0.7022</td>
<td>0.6999</td>
<td>0.6981</td>
<td>0.6965</td>
<td>0.6952</td>
<td>0.6942</td>
<td></td>
</tr>
<tr>
<td>Proportion of managers</td>
<td>0.1113</td>
<td>0.1407</td>
<td>0.1605</td>
<td>0.1679</td>
<td>0.1739</td>
<td>0.1789</td>
<td>0.1830</td>
<td>0.1865</td>
<td>0.1893</td>
<td></td>
</tr>
<tr>
<td>Proportion of supervisors</td>
<td>0.1654</td>
<td>0.1470</td>
<td>0.1346</td>
<td>0.1300</td>
<td>0.1262</td>
<td>0.1231</td>
<td>0.1205</td>
<td>0.1183</td>
<td>0.1165</td>
<td></td>
</tr>
<tr>
<td>Prob γ_m-score is high ability</td>
<td>0.6861</td>
<td>0.6542</td>
<td>0.6412</td>
<td>0.6382</td>
<td>0.6365</td>
<td>0.6357</td>
<td>0.6355</td>
<td>0.6358</td>
<td>0.6363</td>
<td></td>
</tr>
<tr>
<td>Prob γ_s-score is high ability</td>
<td>0.8994</td>
<td>0.8810</td>
<td>0.8775</td>
<td>0.8767</td>
<td>0.8763</td>
<td>0.8761</td>
<td>0.8760</td>
<td>0.8761</td>
<td>0.8763</td>
<td></td>
</tr>
<tr>
<td>Production / supervisory tech.</td>
<td>1.0398</td>
<td>0.9148</td>
<td>0.8320</td>
<td>0.8016</td>
<td>0.7770</td>
<td>0.7566</td>
<td>0.7398</td>
<td>0.7259</td>
<td>0.7143</td>
<td></td>
</tr>
<tr>
<td>Production / managerial tech.</td>
<td>0.3993</td>
<td>0.4985</td>
<td>0.5657</td>
<td>0.5907</td>
<td>0.6114</td>
<td>0.6286</td>
<td>0.6430</td>
<td>0.6550</td>
<td>0.6652</td>
<td></td>
</tr>
<tr>
<td>Total output</td>
<td>1.4391</td>
<td>1.4133</td>
<td>1.3977</td>
<td>1.3925</td>
<td>1.3884</td>
<td>1.3852</td>
<td>1.3828</td>
<td>1.3809</td>
<td>1.3795</td>
<td></td>
</tr>
<tr>
<td>Supervisory share of output</td>
<td>0.7225</td>
<td>0.6473</td>
<td>0.5953</td>
<td>0.5758</td>
<td>0.5596</td>
<td>0.5462</td>
<td>0.5350</td>
<td>0.5256</td>
<td>0.5178</td>
<td></td>
</tr>
<tr>
<td>Managerial share of output</td>
<td>0.2775</td>
<td>0.3527</td>
<td>0.4047</td>
<td>0.4242</td>
<td>0.4404</td>
<td>0.4536</td>
<td>0.4650</td>
<td>0.4744</td>
<td>0.4822</td>
<td></td>
</tr>
<tr>
<td>Relative wages W_w/W_σ</td>
<td>0.7587</td>
<td>0.7323</td>
<td>0.7225</td>
<td>0.7203</td>
<td>0.7190</td>
<td>0.7184</td>
<td>0.7183</td>
<td>0.7185</td>
<td>0.7189</td>
<td></td>
</tr>
<tr>
<td>Supervisor wage at γ_s</td>
<td>1.6711</td>
<td>1.6942</td>
<td>1.7036</td>
<td>1.7058</td>
<td>1.7070</td>
<td>1.7076</td>
<td>1.7077</td>
<td>1.7075</td>
<td>1.7071</td>
<td></td>
</tr>
</tbody>
</table>
\( \alpha^* = .5 \) case: high ability individuals are more scarce
### Changing Test Error: \( a1=5, \; \alpha^*=.5 \)

<table>
<thead>
<tr>
<th>Test score std. dev. ((\sigma))</th>
<th>0.05</th>
<th>0.10</th>
<th>0.20</th>
<th>0.25</th>
<th>0.30</th>
<th>0.35</th>
<th>0.40</th>
<th>0.45</th>
<th>0.50</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test score at dev. ((Q))</strong></td>
<td>1.0012</td>
<td>1.0050</td>
<td>1.0198</td>
<td>1.0308</td>
<td>1.0440</td>
<td>1.0595</td>
<td>1.0770</td>
<td>1.0966</td>
<td>1.1178</td>
</tr>
<tr>
<td><strong>Equilibrium worker wage (W_w)</strong></td>
<td>1.2000</td>
<td>1.1754</td>
<td>1.1103</td>
<td>1.0988</td>
<td>1.0932</td>
<td>1.0910</td>
<td>1.0908</td>
<td>1.0920</td>
<td>1.0939</td>
</tr>
<tr>
<td><strong>Optimal manager cutoff (\gamma_m)</strong></td>
<td>0.5127</td>
<td>0.5169</td>
<td>0.4902</td>
<td>0.4781</td>
<td>0.4680</td>
<td>0.4601</td>
<td>0.4542</td>
<td>0.4502</td>
<td>0.448</td>
</tr>
<tr>
<td><strong>Optimal supervisor cutoff (\gamma_s)</strong></td>
<td>0.7524</td>
<td>0.7465</td>
<td>0.8923</td>
<td>0.9773</td>
<td>1.0657</td>
<td>1.1568</td>
<td>1.2504</td>
<td>1.3463</td>
<td>1.4444</td>
</tr>
<tr>
<td><strong>Proportion low ability</strong></td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
<td>0.6915</td>
</tr>
<tr>
<td><strong>Proportion of workers</strong></td>
<td>0.6957</td>
<td>0.6965</td>
<td>0.6846</td>
<td>0.6786</td>
<td>0.6730</td>
<td>0.6680</td>
<td>0.6634</td>
<td>0.6593</td>
<td>0.6557</td>
</tr>
<tr>
<td><strong>Proportion of managers</strong></td>
<td>0.0781</td>
<td>0.0747</td>
<td>0.1246</td>
<td>0.1499</td>
<td>0.1733</td>
<td>0.1946</td>
<td>0.2138</td>
<td>0.2309</td>
<td>0.2461</td>
</tr>
<tr>
<td><strong>Proportion of supervisors</strong></td>
<td>0.2262</td>
<td>0.2288</td>
<td>0.1908</td>
<td>0.1715</td>
<td>0.1537</td>
<td>0.1374</td>
<td>0.1228</td>
<td>0.1098</td>
<td>0.0982</td>
</tr>
<tr>
<td><strong>Prob (\gamma_m)-score is high ability</strong></td>
<td>0.6000</td>
<td>0.5672</td>
<td>0.4804</td>
<td>0.4651</td>
<td>0.4576</td>
<td>0.4546</td>
<td>0.4544</td>
<td>0.4550</td>
<td>0.4586</td>
</tr>
<tr>
<td><strong>Prob (\gamma_s)-score is high ability</strong></td>
<td>1.0000</td>
<td>0.9932</td>
<td>0.9751</td>
<td>0.9719</td>
<td>0.9703</td>
<td>0.9697</td>
<td>0.9697</td>
<td>0.9700</td>
<td>0.9705</td>
</tr>
<tr>
<td><strong>Production / supervisory tech.</strong></td>
<td>1.1309</td>
<td>1.1436</td>
<td>0.9516</td>
<td>0.8546</td>
<td>0.7649</td>
<td>0.6835</td>
<td>0.6103</td>
<td>0.5451</td>
<td>0.4874</td>
</tr>
<tr>
<td><strong>Production / managerial tech.</strong></td>
<td>0.3027</td>
<td>0.2774</td>
<td>0.4442</td>
<td>0.5301</td>
<td>0.6099</td>
<td>0.6827</td>
<td>0.7485</td>
<td>0.8075</td>
<td>0.8599</td>
</tr>
<tr>
<td><strong>Total output</strong></td>
<td>1.4336</td>
<td>1.4211</td>
<td>1.3958</td>
<td>1.3847</td>
<td>1.3748</td>
<td>1.3662</td>
<td>1.3589</td>
<td>1.3526</td>
<td>1.3473</td>
</tr>
<tr>
<td><strong>Supervisory share of output</strong></td>
<td>0.7889</td>
<td>0.8048</td>
<td>0.6818</td>
<td>0.6172</td>
<td>0.5564</td>
<td>0.5003</td>
<td>0.4492</td>
<td>0.4030</td>
<td>0.3617</td>
</tr>
<tr>
<td><strong>Managerial share of output</strong></td>
<td>0.2111</td>
<td>0.1952</td>
<td>0.3182</td>
<td>0.3828</td>
<td>0.4436</td>
<td>0.4997</td>
<td>0.5508</td>
<td>0.5970</td>
<td>0.6383</td>
</tr>
<tr>
<td><strong>Relative wages (W_w/W_s)</strong></td>
<td>0.6000</td>
<td>0.5798</td>
<td>0.5238</td>
<td>0.5201</td>
<td>0.5160</td>
<td>0.5143</td>
<td>0.5142</td>
<td>0.5151</td>
<td>0.5165</td>
</tr>
<tr>
<td><strong>Supervisor wage at (\gamma_s)</strong></td>
<td>2.0000</td>
<td>2.0274</td>
<td>2.0996</td>
<td>2.1125</td>
<td>2.1187</td>
<td>2.1212</td>
<td>2.1213</td>
<td>2.1200</td>
<td>2.1178</td>
</tr>
</tbody>
</table>

---

**Equilibrium worker wage \(W_w\)**

**Supervisor wage at cutoff**

**Worker, sup. and man. proportions**

**Pr (high ability) at cutoff scores**

- **workers**
- **managers**
- **supervisors**

- **Prob \(\gamma_m\)-score high**
- **Prob \(\gamma_s\)-score high**
References


