5-1-2004

The Effects of Rejected Job Offers on the Costs and Benefits Associated with the Use of Banding Strategies for Employee Selection

Joshua Daniel
Western Kentucky University

Follow this and additional works at: http://digitalcommons.wku.edu/theses
Part of the Industrial and Organizational Psychology Commons

Recommended Citation
http://digitalcommons.wku.edu/theses/546

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
THE EFFECTS OF REJECTED JOB OFFERS ON THE COSTS AND BENEFITS ASSOCIATED WITH THE USE OF BANDING STRATEGIES FOR EMPLOYEE SELECTION

A Thesis
Presented to
The Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment of the Requirements for the Degree Masters of Arts

By
Joshua Daniel

May 2004
THE EFFECTS OF REJECTED JOB OFFERS ON THE COSTS AND BENEFITS ASSOCIATED WITH THE USE OF BANDING STRATEGIES FOR EMPLOYEE SELECTION

Date Recommended: July 2, 2004

Director of Thesis

Dean, Graduate Studies and Research  Date

RECEIVED
JUL - 9 2004

GRADUATE STUDIES & RESEARCH
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Introduction and Review of the Literature</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>16</td>
</tr>
<tr>
<td>Results</td>
<td>18</td>
</tr>
<tr>
<td>Discussion</td>
<td>23</td>
</tr>
<tr>
<td>References</td>
<td>25</td>
</tr>
</tbody>
</table>
THE EFFECTS OF REJECTED JOB OFFERS ON THE COSTS AND BENEFITS ASSOCIATED WITH THE USE OF BANDING STRATEGIES FOR EMPLOYEE SELECTION

Joshua Daniel

Directed by: Reagan D. Brown, Ph.D., Elizabeth L. Shoenfelt, Ph.D., and Jacqueline Pope, Ph.D.

Department of Psychology
Western Kentucky University

Abstract
Among the available selection strategies (e.g., top down selection), sliding bands with minority preference selection was shown to be the most effective at striking a balance between reducing adverse impact with minimal test utility loss. Unfortunately, all previous research into selection strategy effectiveness failed to model job acceptance rates, a variable shown to decrease overall test utility (Murphy, 1986). In this study we compared the utility and adverse impact ratios obtained from strict top down and sliding bands with minority preference selection strategies in which we varied selection ratios, job acceptance rates, and sample sizes. Across all conditions, utility and adverse impact ratios were found to be lower than was demonstrated in previous research which ignored job acceptance rates. Only one of our four hypotheses was supported in this study. We found that differences in adverse impact ratios between top down selection and banding with race preferential selection was reduced when high scoring minority applicants refused offers at a rate higher than high scoring majority applicants. Thus, the benefits that employers expect to see when utilizing the sliding band with minority preference selection strategy are not as great as previously believed.
Introduction and Review of the Literature

There is a well known dilemma that occurs when employers seek to select the best candidates for job openings and create a diverse workforce. This dilemma occurs because minority groups often score lower than majority groups on many of the more popular selection tests. Thus, selecting top scoring candidates often leads to adverse impact; whereas, deliberately selecting more members from the lower scoring group will lead to lower test utility (Campion et al., 2001; Cascio, Outtz, Zedeck, & Goldstein, 1991; Sackett & Roth, 1991; Sackett & Wilk, 1994). Banding is a strategy that is used as a solution to this dilemma.

Since selection tests are not perfect measures and thereby represent a certain amount of unreliability, small differences between scores on a selection test do not necessarily mean that the higher scoring individual will outperform the lower scoring individual on the job. Based on this notion, banding treats individual scores that fall within a specified range as identical. Employers can then deliberately select minority candidates from the score within the band in order to create a diverse workforce and maintain high economic utility of the selection tests (Cascio et al., 1991; Murphy & Myors, 1995; Sackett & Roth, 1991).

The purpose of this study is to challenge the results of studies (e.g., Cascio et al., 1991; Sackett & Roth, 1991) that have concluded that banding is the preferred solution to the dilemma that employers face. In this study we will examine the effects that the use of banding strategies have on utility and adverse impact when job acceptance rates are less than 100%. The results will increase the knowledge about the effectiveness of using such a strategy for achieving a diverse workforce and maintaining high economic utility.
The Dilemma that Faces Employers

Employers have the option to utilize several methods when selecting the best candidates for job openings. Candidates are often selected by their scores on a test or a series of selection procedures (e.g., tests of cognitive or physical ability, personality inventories, weighted application blanks, and situational tests) designed to predict job performance (Cascio et al., 1991; Siskin, 1995). A cutoff score is usually implemented on these predictors as a means of separating more qualified candidates from less qualified candidates. A cutoff score is the point within a distribution of scores that separates scores that pass from those that fail. A passing score meets or exceeds the cutoff; a failing score falls below it.

A dilemma occurs when employers try to reach optimal levels of economic utility while also trying to achieve workforce diversity when selecting job candidates. In most circumstances, reaching one of these goals prohibits achieving the other goal because many minority groups have lower scores on some of the more popular selection tests (Campion et al., 2001; Cascio et al. 1991; Sackett & Roth, 1991; Sackett & Wilk, 1994).

Although under some conditions civil rights laws regulate the achievement of a diverse workforce, organizations also seek to achieve a diverse workforce because it may very well contribute to the overall success. Examples of such reasons include making the organization more attractive to highly qualified minority candidates and customers or creating better solutions to problems through the benefits of heterogeneous groups (Sackett & Wilk, 1994).
Referral Methods

Cascio et al. (1991) explored the statistical implications of six common methods used in setting cutoff scores. These methods, known as referral methods, describe specific ways in which test scores will be used to make selection decisions. Referral methods include the following: strict (i.e., race neutral) top down referral in order of the test scores; within group percentile referral; fixed bands, using random referral within bands; fixed bands, using nonrandom diversity based referral within bands; sliding bands, using random referral within bands; and sliding bands, using nonrandom, diversity based referral within bands.

Strict top down referral is the traditional model for personnel selection. It involves selecting the candidates with the highest scores on a test or a series of selection procedures without regard to group membership (i.e., sex or race) or any measurement error that may be present in these assessments. Applicants are selected in a top down fashion, starting with the highest score, until all positions are filled. Top down referral results in the highest levels of overall performance and economic utility of the test. This method also results in the highest levels of adverse impact of any of the other referral methods. Thus, it is not the preferred method for creating a diverse workforce (Aguinis, Cortina, & Goldberg, 1998; Campion et al., 2001; Cascio et al., 1991; Murphy, 1994; Murphy & Osten, 1995; Siskin, 1995).

In the past, within group percentile referral methods (also known as race norming when done by race) have been used to eliminate adverse impact. This method involves selecting candidates according to their percentile ranking within their group. This method is generally used on selection procedures that yield mean differences between gender or
ethnic groups and result in adverse impact (Sackett & Roth, 1991). If this method is used, it creates the highest selection of minority candidates and eliminates adverse impact but often leads to losses in overall performance and economic utility of the test. This method is often criticized because it ignores obtained raw scores difference between minority and nonminority candidates. A minority candidate could have a score of 70 and be selected over a nonminority candidate with a score of 85 if that candidate’s percentile rank was higher within his or her racial or ethnic group. Furthermore, race norming is explicitly prohibited by the Civil Rights Act of 1991 (Campion et al., 2001; Murphy, 1994; Murphy & Osten, 1995).

Banding

The latter four of these referral methods are forms of banding. Banding has gained popularity because it serves as an alternative to top down selection and helps employers make decisions when faced with the dilemma of creating a diversified workforce and achieving high levels of economic utility from a selection test (Campion et al., 2001). Tests and psychological assessments are never perfect measures and are plagued by certain amounts of unreliability. Banding is based on the notion that small differences in test scores do not necessarily mean that the higher scoring individuals will out perform the lower scoring individual due to the inherent unreliability of the test. Thus, individuals with scores within a specified range of scores are treated as if the scores are identical (Cascio et al., 1991; Murphy & Myors, 1995; Sackett & Roth, 1991).

Statistical data on the error of measurement within the test, the social value, utility, and desire to create workforce diversity influence are variables that influence which candidate will be selected. The various forms of banding differ in the ways that the
test data are incorporated, but usually do not differ in the social and economic goals associated with this method (Cascio et al., 1991). Most of the research on banding has focused on how banding strategies can be used to reduce adverse impact and minimize losses in utility. Evidence has been found that supports this notion (Cascio et al., 1991; Report of the Scientific Affairs Committee, 1994; Sackett & Roth, 1991).

**Social Value of Banding**

Banding is commonly used to reduce the levels of adverse impact that may result from a typical selection test. If two ethnic groups differ in their test score means, relatively few members from the lower scoring group will be selected on the basis of their scores. Banding is a way of selecting candidates that can result in a greater number of minority candidates being selected and lowers the probability of adverse impact (Murphy & Osten, 1995). Campion et al. (2001) pointed out that the use of banding has increased since the passage of the 1991 Civil Rights Act. Banding is not a new concept, however. Most organizations utilize some form of banding, whether it is a pass/fail cutoff, a descriptive classification of candidates based on some scored criterion, or a letter grade assigned to a group of scores, which is done in the academic setting.

**Statistical Rationale for Banding**

The size, or width, of a band is determined by four variables: the reliability of the test, the standard deviation of the test, a constant of 1.41 (i.e., the square root of two), and how confident one wants to be that two scores are reliably different from each other. Lower reliabilities lead to larger bandwidths. Increased confidence (e.g., 99% versus 95%) and/or larger standard deviation of test scores also lead to larger bandwidths. The confidence value should be one tailed (e.g., 1.65 for 95% confidence) because all scores
are compared against the highest score in the band. Scores outside the band are reliably lower than the highest score in the band (assuming sliding bands). The equation for bandwidth is given below.

\[
\text{Bandwidth} = C \times SDx \times \sqrt{(1-r_{xx})} \times \sqrt{2}
\]  

(1)

Where

- \( C \) is the confidence value (e.g., 1.65 for 1-tailed)
- \( SDx \) is the standard deviation of the test
- \( r_{xx} \) is the reliability of the test

Once bandwidth is established, the first band is formed by subtracting the bandwidth from the highest score. All scores in that range are in the band. All scores outside the range are outside the band. For example, if bandwidth is 10 points and the highest score is 98, then the band includes all scores from 98 to 88. A score of 87 is outside the band and is considered, given the error of measurement on the test, to be reliably lower than 98. A score of 88, however, is considered to be the same score as 98, given the amount of measurement error in the test.

**Bandwidth Issues**

Statistically, bandwidth varies according to the reliability of the test, the standard deviation of the test, and the confidence interval \( C \) placed on the band. The confidence interval is the only area that the employer has statistical control over when creating the band. A higher confidence interval creates a wider band and incorporates more members from a lower scoring group. It is not uncommon to see a 30% increase in bandwidth.
when moving from a 95% to 99% confidence interval, or a 50% increase in bandwidth when moving from a 68% to 95% confidence interval. A higher confidence interval creates greater assurance that candidates outside of the band reliably differ from the highest score within the band (Murphy, 2001).

Campion et al. (2001) suggested that further research needs to be conducted to determine other factors that may be relevant when creating a band. These areas include the nature of the job, the legal and political environment pertaining to a diverse workforce, and the business necessity for a diverse workforce. For instance, they stated that bandwidth should be narrower for jobs that involve high risks for incorrect decisions (e.g., police officer) and should be wider for jobs that are low in risk. Narrower bands would make it less probable that lower scoring candidates to be included and lower the risk of selecting a candidate with a lower true score that would perform worse in the actual job setting.

*Fixed vs. Sliding Bands*

There are two forms of banding: fixed or sliding. Both reference the same bandwidth for a given standard error of measurement and confidence interval. Both types also reference the highest observed score to create the upper point of the band. However, they differ in the constancy of the candidates within the band (Murphy & Osten, 1995). In a fixed band, the location of the band (e.g., scores ranging from 86 to 95) remains constant. An example of fixed bands exists in the college grading system. This system groups a specified range of scores and assigns a letter grade to scores within that range. Thus scores of 80-89 might be considered equal and all translate into a “B” letter grade, and scores of 89 and 90 might be treated as different and translate into “B” and “A” letter
grades (Murphy, 1994). Candidates within fixed bands are selected through random or nonrandom procedures until the band is exhausted of candidates. If more job openings remain after the top band has been depleted of candidates, candidates from a second band are chosen (Cascio et al., 1991).

Sliding bands differ from fixed bands because their location changes every time the top scorer is chosen, regardless of whether or not other candidates remain within the band. If the top scorer is chosen and other candidates remain within the same band, the next highest remaining candidate is then used as a reference point for the formation of a new band. Thus, the highest scorer outside of the original band (which was previously considered different from the highest scores) may now be included into the top band and considered statistically equal to the remaining scores within that band. Sliding bands provide more opportunities for selection of minority candidates than do fixed bands because they move through the distribution more quickly (Murphy & Osten, 1995). They have been shown to produce minority selection rates comparable to the proportion of minorities within the applicant pool (Cascio et al., 1991; Sackett & Roth, 1991).

Selection Within Bands

Once a band is formed, selection of candidates within the band can be random or nonrandom. Random selection allows for each candidate within the band to have an equal chance of being selected and is used with the rationale that all candidates within the band are not reliably different from each other and therefore performance differences should not exist. Nonrandom selection occurs when applicants are chosen on some criteria other than their observed score on the test in question (e.g., diversity, job experience, training, seniority, score on another test). Use of minority preference as a tie breaker is done to
reduce levels of adverse impact and create optimum economic utility for the test or selection procedure (Cascio et al., 1991; Cascio, Goldstein, Outtz. & Zedeck, 1995; Murphy & Osten, 1995).

Sackett and Roth (1991) reported that random selection within bands creates the same rate of minority hiring as strict top down selection. Banding can only be expected to increase minority hiring when a nonrandom procedure with minority preference is used. Specifically, sliding bands with minority preference selection creates the highest rate of minority hiring of the banding approaches (Cascio et al., 1991; Sackett & Roth 1991). Cascio et al. (1991) reported a 47% minority hiring rate when sliding bands were used with minority preference, which did not result in adverse impact (i.e., the 4/5th rule was not violated). In comparison, the strict top down and the fixed band with minority preference approaches resulted in 20% minority hiring rates, which resulted in adverse impact. The sliding band with random selection resulted in 14% minority hiring rate, which also resulted in adverse impact. Thus, sliding bands with diversity based referral may be the preferred strategy for employers desiring both high utility and diversity from their selection tests.

Selection within the band is also affected by the characteristics of the applicant pool. The proportion of applicants from the lower scoring group, group differences in mean test scores, and group differences in the standard deviation of the test scores all affect the outcome that a banding strategy will produce. The proportion of applicants from the lower scoring group affects the total number of candidates who will be selected from that group. If there is a low proportion, fewer candidates will be selected from that group. The inverse is true for a high proportion of candidates. Likewise, the differences
in mean test scores between two groups affect the selection likelihood of a candidate from the lower scoring group. Larger differences between mean scores of the two groups result in less probability of the lower scoring group being selected. Smaller differences result in a higher probability of selection. The difference in the standard deviation of the two groups will also affect the selection outcome. If the standard deviation is large within either group, it will increase the number of applicants selected from the lower scoring group (Murphy & Osten, 1995).

The selection ratio can also affect the proportion of candidates chosen from the lower scoring group when bands are used. Lower scoring candidates benefit the least if the selection ratio is very high or low and the reliability of the test is very high. The number of candidates chosen from the lower scoring group is most beneficial (i.e., candidates from the lower scoring group have a higher probability of selection) when the selection ratio is moderate and the reliability of the test is low (Murphy, 1994; Murphy & Osten, 1995; Sackett & Roth, 1991). Therefore, the same banding approach may be applied to different situations and have different results as a function of dissimilar selection ratios (Report of the Academic Affairs Committee, 1994).  

*Test Utility and Banding*

Test utility refers to how useful a test is to the organization. Sackett and Roth (1991) suggested that banding would result in overall losses in selection utility in comparison to top down referral. Specifically, wider bands result in the selection of lower test score averages. Therefore, narrower bands result in higher utility than wider bands. Utility of the band is also affected by how the selection occurs within the band (Murphy, 1994). According to Zedeck, Outtz, Cascio, & Goldstein (1991), it is reasonable for an
organization to suffer some loss in utility in order to improve the diversification of their workforce. Banding is one alternative to top down referral that allows for minimal losses of utility in an effort to create diversity. Zedeck et al. (1991) referenced the Cascio et al. (1991) data set to support their statement. Cascio et al. reported a mean score of 88.79 for sliding bands with minority preference in comparison to a score of 89.81 for a race blind, top down approach. The similarity of mean scores illustrates the minimal loss in utility when the banding strategy is used in comparison to the top down approach.

Criticisms of Banding

Kehoe and Tenopyr (1994) illustrated a flaw that is evident within banding. Given that candidates within the band are considered equal due to measurement within the test, retest data should show different scores for candidates within the band. However, their retest data showed that given typical levels of reliability and confidence (i.e., 95%) candidates at the top of the band can be 25 times more likely to score higher than candidates at the bottom of the band. Using criterion data, they also showed similar data from which they concluded that scorers at the top of the band could be twice as likely to outperform scorers at the bottom of the band while on the job. However, contrary to the findings of Kehoe and Tenopyr, Siskin (1995) presented a mathematical model that illustrated that in most situations, the top scorer in the band is only slightly more likely to outperform the bottom scorer.

Aguinis et al. (1998) also agreed with the postulate that performance levels differ within a band. They reported that applicants that fall within the same band can easily have different performance levels. Along the same line, Schmidt (1991) reported that the band must then be formed on the bases of a regression line that predicts job performance.
from test scores. If a band is formed based on this principle, tests with a validity of .50 will include 97% of all candidates at the 95% confidence interval. Essentially, all candidates would be placed in the same band, rendering banding virtually useless (Campion et al., 2001).

Another criticism of banding concerns the motives that people have when using an unreliable test. Since greater unreliability within a test results in a wider bandwidth, people could purposely use less reliable tests to create a wider band. However, this approach leads to a less valid test and is considered unethical, which reduces the likelihood that it will occur (Cascio et al., 1995). Cascio et al. noted that this approach could occur for top down selection as well, but critics only make the point in reference to banding. Regardless of the validity of these criticisms, banding still offers a way for organizations to meet the dual goals of high utility and workforce diversity.

**Legality of Banding**

The Civil Rights Act of 1991 illustrated the controversy surrounding the issue of minority preference in selection procedures by prohibiting any group based score adjustments. Thus, any banding procedure that is used for the reason of minority preference could be successfully challenged in court (Report of the Scientific Affairs Committee, 1994). The statistical rationale behind banding has been upheld in various court cases (Bridgeport Guardians v. City of Bridgeport, 1991; Office for Justice v. Civil Service Commission, 1992). Thus, the psychometric and statistical rationale for placing scores in a band is beyond legal dispute. It is nonrandom selection of applicants from within a band that rests on questionable legal footing. However, these court cases have not supported banding when it is used with minority preference. Banding allows for
choices to be made on some criteria other than test scores such as job experience, seniority, prior job performance, licensing, etc. It was recommended that professional conduct, education, and training/experience be used as the secondary criteria in the *Officers v. CSC* (1992) case (Gutman & Christiansen, 1997; Campion et al., 2001). However, when group membership (i.e., race and sex) is the variable used to break ties within the band, its legality is questionable. The most recent and relevant rulings from the Supreme Court (the University of Michigan cases: *Gratz et al. vs. Bollinger et al.*, 2003; *Grutter vs. Bollinger et al.*, 2003) have failed to unambiguously answer the question as to whether group membership can be used to select applicants within bands. Thus, the door to using banding to achieve the goals of high utility and workforce diversity has been left open.

*The Present Study*

Most utility estimates, regardless of banding issues, make one assumption: job offers will be accepted 100% of the time. Murphy (1986) found that utility estimates overestimate utility gains by 30%-80% when this assumption is not met. To this point, no one has examined the utility of banding as well as the diversity increases of banding when the assumption of 100% job offer acceptance is not met. The present study will do so.

Hypothesis 1a: Utility differences between top down selection and sliding bands with race preferential selection will be reduced when some of the top scoring applicants refuse or are not offered the job.

Murphy (1986) reported that more than 35% of job offers are turned down in the
technical and engineering fields and more than 25% of jobs are turned down in non-technical fields. Candidates who no longer desire a specific job can be expected to refuse that job offer. This refusal can occur because the candidate has already accepted or plans to accept an alternative job, has plans to pursue other opportunities, or because of personal reasons. It is unknown whether the refusal of job offers will affect the gains in diversity offered by sliding bands with minority preference versus strict top down selection.

Hypothesis 1b: Adverse impact differences between top down selection and sliding bands with race preferential selection will be reduced when some of the top scoring applicants refuse or are not offered the job.

Intuitively, the highest scoring minorities can also be expected to turn down the job at a higher rate than the higher scoring majorities. This follows the logic that high scoring minorities are more likely to be in demand by various organizations than their high scoring majority counterparts and are therefore more likely to turn down job offers because of their available options. Thus, another hypothesis was created to illustrate what would happen to utility and adverse impact if minorities turned down the job at a higher rate than minorities.

Hypothesis 2a: Utility differences between top down selection and sliding bands with race preferential selection will be reduced when high scoring minority applicants refuse offers at a rate higher than high scoring majority applicants.

Hypothesis 2b: Adverse impact differences between top down selection and sliding bands with race preferential selection will be reduced when high scoring minority
applicants refuse offers at a rate higher than high scoring majority applicants.
Method

Data Generation

Normally distributed test score data were generated using a program created in the C programming language. The simulated data were used to represent applicant test scores for both minority and majority groups. The majority group applicants were randomly generated in a population with a mean score of 100 and a standard deviation of 15. The minority group applicants were randomly generated in a population with a mean score of 85 and a standard deviation of 15. Thus, minority group members' scores were modeled to average one standard deviation lower than majority group members (Sackett & Roth, 1991). Scores were then rounded to their nearest whole number. Simulated applicants were then selected from the total applicant pool through use of strict top down and sliding bands with race preferential selection strategies.

We simulated field conditions in which each applicant pool had a fixed percentage of minority and majority applicants (50% for each group). We selected a condition of equal representation in the applicant pool because Sackett and Roth's (1991) study showed the best results for banding (in terms of highest adverse impact ratios) when the applicant pool was evenly divided. The within group standard deviation (15), reliability (.9), and confidence value (95%, one tailed) were then used to compute the bandwidth, which was calculated to be 11.07 points. The two selection methods (strict top down and sliding band with minority preference) were then applied to the total simulated applicant pool. The following variables were used to model applicant scores:

1. Total Selection ratio (across both groups): 10% and 30%.
2. Total sample size: 100 and 600.
3. Percent of job offers that would be accepted if the applicant was offered the job: 100% for both groups, 90% for both groups, and 90% for majority group / 80% for minority group.

For those conditions in which some of the job offers would not be accepted, the dichotomous variable of rejecting/accepting the job offer was generated in a manner such that it correlated approximately .5 (correlations range from .45 to .55) with applicant test score; that is, higher scoring applicants (presumably with more job opportunities) were more likely to reject the job offer. One thousand simulations were run for each condition.

Analyses

For all analyses, the dependent variables were the mean test score of all selected applicants and the adverse impact ratios for each of the selection strategies. Each simulation in each condition yielded four dependent variables: mean test score with top down selection, mean test score with sliding bands with race preferential selection, adverse impact ratio with top down selection, and adverse impact ratio with sliding bands with race preferential selection. We computed the difference between mean test scores and adverse impact ratios which reduced the number of dependent variables from four to two for simulation for each condition. Thus, each dependent variable reflected the difference in either utility or adverse impact ratios that would be obtained from the two selection strategies. Mean differences between conditions were examined with independent samples \( t \) tests.
Results

Tables 1 and 2 list the mean (averaged across 1000 simulations) adverse impact ratios and mean test scores of candidates selected for the job through the use of strict top down and sliding band with minority preference selection strategies. Table 1 illustrates results when the sample size is 100 and Table 2 illustrates results when the sample size is 600. The tables summarize data from the 10% and 30% selection ratios crossed with the 100%, 90%, and 90% (majority) / 80% (minority) job acceptance rates.

The results show a number of findings, some expected, others unexpected. Results were nearly identical for both the 100 and 600 sample size conditions; however, there were a few more deviant results when the sample size was 100. As expected, strict top down selection produced higher mean test scores. Utility (as indexed by mean score of applicants hired) is lower for higher selection ratios. Within each condition, adverse impact ratios were better and utilities were worse for banding based selection versus top down selection. Both utility and adverse impact ratios are lower when high scoring applicants decline the job offer. Finally, when minorities turn down the job at a rate greater than majorities and banding is used to make selection decisions, utility increases and adverse impact ratios decrease.

A couple of unexpected results were also observed. First, in a number of instances, banding based selection resulted in lower adverse impact ratios at higher selection ratios. This finding is contrary to existing research (Sackett & Wilk, 1994), which shows higher (i.e., better) adverse impact ratios for higher selection ratios. Second, adverse impact ratios under top down selection actually improved when high scoring minority candidates refused the job offer at rates higher than majority candidates.
Table 1

*Mean Test Scores of Selected Applicants (Utility) and Adverse Impact when N = 100*

<table>
<thead>
<tr>
<th>Total</th>
<th>Mean Score (Utility)</th>
<th>Mean Score (Utility)</th>
<th>Adverse Impact</th>
<th>Adverse Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>Top Down</td>
<td>Banding</td>
<td>Top Down</td>
<td>Banding</td>
</tr>
</tbody>
</table>

100% Job Acceptance Rates

<table>
<thead>
<tr>
<th>10</th>
<th>120.8</th>
<th>119.1</th>
<th>16.1</th>
<th>81.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>111.3</td>
<td>110.1</td>
<td>29.1</td>
<td>79.2</td>
</tr>
</tbody>
</table>

90% Job Acceptance Rates

<table>
<thead>
<tr>
<th>10</th>
<th>114.3</th>
<th>112.6</th>
<th>15.6</th>
<th>79.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>105.0</td>
<td>103.6</td>
<td>25.8</td>
<td>75.8</td>
</tr>
</tbody>
</table>

90% Majority and 80% Minority Job Acceptance Rates

<table>
<thead>
<tr>
<th>10</th>
<th>114.8</th>
<th>113.3</th>
<th>20.2</th>
<th>79.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>105.1</td>
<td>103.9</td>
<td>24.2</td>
<td>66.4</td>
</tr>
</tbody>
</table>

*Note.* All entries are mean scores across 1000 simulations. When jobs were accepted 90%, majority and minority group members turned down the job at a .49 correlation with their score. When jobs were accepted by 90% of majority group member and 80% of minority group members, majority group members turned down the job at a .48 correlation and minority group members turned down the job at a .50 correlation with their test score.
Table 2

*Mean Test Scores of Selected Applicants (Utility) and Adverse Impact when N = 600*

<table>
<thead>
<tr>
<th>Total Selection Ratio</th>
<th>Top Down Banding</th>
<th>Top Down Banding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean Score</strong> (Utility)</td>
<td><strong>Mean Score</strong> (Utility)</td>
<td><strong>Adverse Impact</strong></td>
</tr>
<tr>
<td>100% Job Acceptance Rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>121.3</td>
<td>119.7</td>
</tr>
<tr>
<td>30</td>
<td>111.5</td>
<td>110.3</td>
</tr>
<tr>
<td>90% Job Acceptance Rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>115.3</td>
<td>113.7</td>
</tr>
<tr>
<td>30</td>
<td>105.5</td>
<td>104.3</td>
</tr>
<tr>
<td>90% Majority and 80% Minority Job Acceptance Rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>115.5</td>
<td>114.0</td>
</tr>
<tr>
<td>30</td>
<td>105.6</td>
<td>104.3</td>
</tr>
</tbody>
</table>

*Note.* All entries are mean scores across 1000 simulations. When jobs were accepted 90%, majority and minority group members turned down the job at a .46 correlation with their score. When jobs were accepted by 90% of majority group member and 80% of minority group members, majority group members turned down the job at a .46 correlation and minority group members turned down the job at a .51 correlation with their test score.
Decreasing the representation of minority candidates relative to majority candidates should lower adverse impact (i.e., make it worse). Given that these results reflect the average across 1000 simulations, sampling error is an unlikely cause.

With regard to our hypotheses, the results were mixed. Hypothesis 1a stated that the utility difference between selection strategies (top down versus sliding band with minority preference) would narrow when some high scoring applicants (10% of minority and 10% of majority) refused the job offer. To test this (and all subsequent hypotheses), we conducted independent sample t tests between the 100% acceptance simulations and the simulations in which some of the applicants refused the job offer. For each hypothesis, four comparisons were made: 100% job acceptance compared with less than 100% job acceptance at each corresponding selection ratio (10% or 30%) for both sample sizes (100 or 600). Unfortunately, Hypothesis 1a was only supported in one of the four comparisons (n = 600, 10% selection ratio), \( t(1998) = 2.74, p < .05 \). Given only one significant result out of four and given that in one of the other three comparisons, the mean difference was large enough to be significant in the wrong direction (i.e., the utility difference actually increased), we conclude that Hypothesis 1a was not supported.

Results were similarly mixed for Hypothesis 1b, which stated that adverse impact difference between the two strategies would narrow when higher scoring applicants from both groups refused the job offer. Independent samples t tests revealed significant mean differences in only one of the four comparisons (n = 600, 30% selection ratio), \( t(1998) = 2.56, p < .05 \). Thus, we conclude Hypothesis 1b was not supported. In short, differences in utility and adverse impact between the alternative selection strategies (top down versus
bANDING) do not change when higher scoring applicants from both groups refuse the job offer in equal proportions.

Hypotheses 2a and 2b, similar to Hypotheses 1a and 1b, stated that utility and adverse impact differences between the selection strategies would narrow when high scoring minority group members refused the job at higher rates than majority group members. Similar to the results found for Hypothesis 1a, the results for Hypothesis 2a were mixed, with significant differences for two of the four comparisons (n = 100, 10% selection ratio and n = 600, 10% selection ratio), \( t(1998) = 5.68, p < .05 \) and \( t(1998) = 14.74, p < .05 \) respectively. Thus, we conclude that Hypothesis 2a was not supported. Hypothesis 2b, however, was supported with significant results in all four comparisons, \( t(1998) = 2.84, p < .05 \) for the least significant comparison. Thus, we conclude that although utility differences between the selection strategies remain constant the differences in adverse impact between the two selection strategies diminish (banding becomes profoundly less effective at hiring minorities) when minorities refuse the job offer more often than majority group members.
Discussion

Previous studies (e.g., Cascio, 1991; Sackett and Roth, 1991) have established the costs (lower test utility) and benefits (reduced adverse impact problems) associated with the use of sliding bands with minority preference as compared to strict top down selection. Unfortunately, all job offers are not accepted by applicants, reducing the utility of testing (Murphy, 1986). The effects of rejected job offers on banding based selection as compared with top down selection strategies have not been examined to date. Our study attempted to identify and model the true costs and benefits of banding (utility and adverse impact) in a simulation that resembles the reality of selection (i.e., with rejected job offers). We found that on an absolute level, both utility and adverse impact ratios were lowered for both top down and sliding band selection when applicants refused the job offer, but the differences between the selection strategies in terms of utility and adverse impact ratios were unchanged as long as members of both groups refused the job in the same proportion. When minority group applicants refused the job offer at a rate greater than the majority group (20% versus 10% job offer rejection rates, respectively), utility differences between the two selection strategies were again unchanged, but the difference in adverse impact ratios (again, between the available selection strategies) were reduced. Thus, the benefit of using banding as opposed to top down selection, in terms of more favorable adverse impact ratios, was smaller in magnitude when minority group members refused job offers at a rate greater than the majority group.

The only limitations with our study concerned the realism of the assumptions used in our model. Specifically, if minority group members refuse job offers at a rate equal to or less than majority group members, then Hypothesis 2b (minority group members
refuse the job at a higher rate than majorities), although supported in our analyses, would be irrelevant. Alternatively, if minority group members refuse job offers at a rate greater than majority group members, then the benefits of banding in terms of more desirable adverse impact ratios over top down selection would be further reduced.

Future research regarding the costs and benefits of banding should explore two avenues. First, the previously mentioned job offer rejection rates should be modeled to include even more skewed ratios (e.g., 0% rejection for majority group versus 20% for minority group). Second, the relative size of the majority and minority groups in the applicant pool should also be modeled to ratios other than the 50/50 split used in this study. We employed an equal representation condition because Sackett and Roth’s (1991) analysis demonstrated it to have the best results (i.e., biggest benefit) in terms of adverse impact. Other, more skewed ratios might accelerate the decline in benefits in terms of adverse impact for banding based selection strategies.

The use of sliding bands with minority preference for selection offers clear benefits to the employer at little cost to test utility (Sackett & Roth, 1991). The research presented here indicates that the benefits might not be as great as previously thought. Furthermore, the legality of using a minority preference rule to break the ties within bands is questionable given the most recent Supreme Court rulings. Future research or future litigation might invalidate any benefit associated with banding and render (or expose) the selection strategy to be functionally impotent.
References


Bridgeport Guardians. Inc. v. City of Bridgeport (CA2, 1991) 933 F.2d 1140.


Officers for Justice v. The Civil Service Commission, 979 F.2d 721 (9th Cir. 1992).


