

**THE NEED FOR LEGISLATIVE OR JUDICIAL CLARITY
ON THE FOUR-FIFTHS RULE AND HOW EMPLOYERS IN
THE SIXTH CIRCUIT CAN SURVIVE THE AMBIGUITY**

SCOTT W. MCKINLEY*

I. INTRODUCTION

The Equal Employment Opportunity Commission (EEOC) provides uniform guidelines for employers to follow in administering promotional examinations.¹ One such guideline is known as the “four-fifths rule.”² The four-fifths rule essentially states that a selection rate for any protected class under Title VII which is greater than four-fifths of the rate of the majority group “will generally not be regarded by Federal enforcement agencies as evidence of adverse impact.”³

In a recent Sixth Circuit decision, *Isabel v. City of Memphis*,⁴ the employer hired an Industrial Organizational Psychologist, with the agreement of the union, to develop a promotional examination that would be nondiscriminatory and which would pass the EEOC’s four-fifths rule.⁵ Although the promotional examination passed the four-fifths rule, the court looked to alternative statistical analyses—the *t*-test and *z*-score—and found the promotional examination to be unlawfully discriminatory.⁶

As a result of this case, employers in the Sixth Circuit, like those in several other circuits, can no longer rely on the EEOC’s four-fifths rule to ensure that their tests are nondiscriminatory. Several circuit and lower courts are looking to alternative analyses and are giving little, if any, deference to the four-fifths rule.⁷ This Comment is a call for legislative or judicial clarity on the four-fifths rule. If the EEOC’s four-fifths rule is no

Copyright © 2008, Scott W. McKinley.

*J.D. Candidate, May 2009, from Capital University Law School. I would like to thank Professor Floyd Weatherspoon, Rick Leslie, and Mēgan Elasky for their inspiration and assistance.

¹ 29 C.F.R. §§ 1607.1–.16 (2008).

² *Id.* § 1607.4(D).

³ *Id.*

⁴ 404 F.3d 404 (6th Cir. 2005).

⁵ *Id.* at 408–09.

⁶ *Id.* at 409.

⁷ *See infra* Part III.B.

longer a reliable guideline for employers to turn to in determining whether their promotional examinations are discriminatory, then the four-fifths rule should be replaced with a more accurate guideline. On the other hand, if the four-fifths rule is determined to be the most accurate guideline in establishing adverse impact, the courts should afford greater deference to the rule. Either way, it is clear that the currently ambiguous status of the four-fifths rule has a negative impact on employers and employees, and it undermines the purpose of Title VII of the Civil Rights Act of 1964.

II. BACKGROUND

A. Title VII's Prohibition of Employment Discrimination in Promotional Examinations

Title VII of the Civil Rights Act of 1964 states that “[i]t shall be an unlawful employment practice for an employer . . . to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment because of such individual’s race, color, religion, sex, or national origin”⁸ Furthermore, in regards to the use of test scores, the Act states the following:

It shall be an unlawful employment practice for a respondent, in connection with the selection or referral of applicants or candidates for employment or promotion, to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex, or national origin.⁹

In enacting Title VII, Congress intended to “achieve equality of employment opportunities and remove barriers that have operated in the past to favor an identifiable group of white employees over other employees.”¹⁰

Federal courts applying Title VII have established that discrimination in employment promotion decisions falls within the statute’s general proscriptions.¹¹ Therefore, employers covered under Title VII must ensure

⁸ 42 U.S.C. § 2000e-2(a) (2000).

⁹ *Id.* § 2000e-2(l).

¹⁰ *Griggs v. Duke Power Co.*, 401 U.S. 424, 429–30 (1971).

¹¹ 48 AM. JUR. 3D *PROOF OF FACTS* 75, 86–87 (1998).

that their promotional examinations do not have a discriminatory effect on a protected class of employees.¹²

B. The Origin of the Disparate Impact Claim Under Title VII

Early Title VII cases frequently targeted facially discriminatory policies “such as segregated seniority rosters or blanket racial exclusions from particular positions.”¹³ However, in the landmark Supreme Court decision of *Griggs v. Duke Power Co.*,¹⁴ the Court, for the first time, accepted the disparate impact claim whereby an individual protected under Title VII, who was disparately impacted by a facially neutral employment policy, could seek relief without having to prove intentional discrimination by the employer.¹⁵ Prior to the *Griggs* decision, employers could attempt to maintain discriminatory employment practices by implementing facially neutral practices that had a discriminatory effect.¹⁶ The Court clarified that under Title VII, “practices, procedures, or tests neutral on their face, and even neutral in terms of intent, cannot be maintained if they operate to ‘freeze’ the status quo of prior discriminatory employment practices.”¹⁷ This holding opened the doors for impact to be used to prove discrimination.

1. Proving a Disparate Impact Claim

The first step in proving a disparate impact claim under Title VII requires the complaining party or class to make out a prima facie case of

¹² *Id.* at 89.

¹³ Allan G. King, “Gross Statistical Disparities” as Evidence of a Pattern and Practice of Discrimination: Statistical Versus Legal Significance, 22 LAB. LAW. 271, 274 (2007) (citing *Franks v. Bowman Transp. Corp.*, 495 F.2d 398, 409–11 (5th Cir. 1974), *rev’d*, 424 U.S. 747 (1976)).

¹⁴ 401 U.S. 424 (1971).

¹⁵ *Id.* at 431 (“The Act proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation.”).

¹⁶ For example, in the *Griggs* case, the employer, who had engaged in intentional discriminatory conduct prior to the enactment of Title VII, implemented a policy that conditioned employment transfers on passing an intelligence test. *Id.* at 428. The test had a discriminatory effect on African-American employees and applicants; however, because the test was facially neutral, the employer’s subjective discriminatory intent could not be proved. *See id.* at 432.

¹⁷ *Id.* at 430.

discrimination that shows “that the tests in question select applicants for hire or promotion in a racial pattern significantly different from that of the pool of applicants.”¹⁸ Once a prima facie case of discrimination is made, the burden then shifts to the employer to show “that any given requirement [has] . . . a manifest relationship to the employment in question.”¹⁹ The *Albemarle* Court clarified the job-relatedness prong as to require “professionally acceptable methods, to be ‘predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated.’”²⁰

The EEOC has issued Guidelines for employers to use to validate whether their tests are job related.²¹ The EEOC’s Guidelines are consistent with professional standards²² and constitute “[t]he administrative interpretation of the Act by the enforcing agency,” and consequently they are ‘entitled to great deference.’”²³ The EEOC provides the following three methods that can be used to validate an employment test: (1) criterion-related validation;²⁴ (2) content validity;²⁵ and, (3) construct validity.²⁶

If the employer is able to prove that its tests are valid and, therefore, job related, the complaining party or class may still be able to prove a disparate impact claim by showing that “other tests or selection devices, without a similarly undesirable racial effect, would also serve the employer’s legitimate interest in ‘efficient and trustworthy

¹⁸ *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975) (citing *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 802 (1973)).

¹⁹ *Griggs*, 401 U.S. at 432.

²⁰ *Albemarle Paper Co.*, 422 U.S. at 431 (citing 29 C.F.R. § 1607.4(c) (1975)).

²¹ See 29 C.F.R. §§ 1607.1–.16 (2008).

²² See *id.* § 1607.5(C) (“Guidelines are consistent with . . . Standards for Educational and Psychological Tests prepared by a joint committee of the American Psychological Association, the American Educational Research Association, the National Council on Measurement in Education, and standard textbooks and journals in the field of personnel selection.”).

²³ *Albemarle Paper Co.*, 422 U.S. at 431 (quoting *Griggs*, 401 U.S. at 433–34).

²⁴ 29 C.F.R. § 1607.14(B).

²⁵ *Id.* § 1607.14(C).

²⁶ *Id.* § 1607.14(D).

workmanship.”²⁷ By making this showing, the complaining party or class would strengthen its evidence that the employer’s tests were used as a “pretext” for discrimination.²⁸ This third prong was subsequently weakened in *Wards Cove Packing Co., Inc. v. Atonio*,²⁹ where the Court held that great deference should be afforded the employer’s decision in adopting alternative employment practices, and the “judiciary should proceed with care before mandating that an employer must adopt a plaintiff’s alternative selection or hiring practice in response to a Title VII suit.”³⁰ However, Congress responded to the added deference afforded to employers in the *Wards Cove* holding by codifying the disparate impact claim and curtailing the deference afforded to employers.³¹ A plaintiff’s use of statistics to prove a disparate impact claim, therefore, remains crucial to the success of the suit.

2. Using Statistics to Prove Disparate Impact

Until the mid 1970s, the statistical analysis in discrimination cases was “largely intuitive.”³² However, in 1977 the Supreme Court first used statistics to prove discrimination in a jury selection case by calculating standard deviations.³³ That same year, the Supreme Court also allowed a plaintiff to make out a prima facie case of employment discrimination under Title VII using statistics to show disparate impact.³⁴ The Court stated that “statistics are not irrefutable, they come in infinite variety and, like any other kind of evidence, they may be rebutted. In short, their usefulness depends on all the surrounding circumstances.”³⁵

Affirming the use of statistics for proving a prima facie case of discrimination under Title VII, the Supreme Court refined the use of statistics by drawing the sample from the actual labor market, rather than

²⁷ *Albemarle Paper Co.*, 422 U.S. at 425 (quoting *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 801 (1973)).

²⁸ *See id.* at 425; *McDonnell Douglas Corp.*, 411 U.S. at 804–05.

²⁹ 490 U.S. 642 (1989).

³⁰ *Id.* at 661.

³¹ Civil Rights Act of 1991, 42 U.S.C. § 2000e-2(k) (2000).

³² *Mitchell v. Rose*, 570 F.2d 129, 133 n.4 (6th Cir. 1978).

³³ *Castaneda v. Partida*, 430 U.S. 482, 496–97 n.17 (1977).

³⁴ *Int’l Bhd. of Teamsters v. United States*, 431 U.S. 324, 339 (1977).

³⁵ *Id.* at 340.

from the population at large.³⁶ With the use of statistics becoming more prevalent in proving disparate impact claims under Title VII, the EEOC provided Uniform Guidelines on Employee Selection Procedures, which included the four-fifths rule.³⁷

C. *The EEOC's Four-Fifths Rule*

Under the Uniform Guidelines on Employee Selection Procedures adopted by the EEOC, “a selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact”³⁸ This is known as the “four-fifths rule.”³⁹ If an employer implements a promotional examination and sets the cutoff score at a level that results in a selection rate for any race, sex, or ethnic group which is greater than four-fifths of the rate for the group with the highest rate, the promotional examination will “generally not be regarded by Federal enforcement agencies as evidence of adverse impact.”⁴⁰

Employers look to the EEOC's Guidelines in administering promotional examinations and often rely on the four-fifths rule as an indicator as to whether their examinations have an adverse impact.⁴¹ However, as this Comment points out, after the *Isabel* decision employers in the Sixth Circuit, like several other circuits, can no longer rely on the EEOC's four-fifths rule. Several circuit courts are looking to alternative statistical analyses in deciding whether a particular employment practice has an adverse impact.⁴² In fact, several circuit courts—especially the

³⁶ *Hazelwood Sch. Dist. v. United States*, 433 U.S. 299, 308 (1977).

³⁷ Adverse Impact and the Four-Fifths Rule, 29 C.F.R. § 1607.4(D) (2008).

³⁸ *Id.*

³⁹ *Id.*

⁴⁰ *Id.*

⁴¹ See Scott B. Morris & Russell Lobsenz, *Significance Tests and Confidence Intervals for the Adverse Impact Ratio*, 13th Annual Society for Industrial and Organizational Psychology Conference 5 (1998) (“The fact that the four-fifths rule is relatively straightforward and easy to implement has resulted in its becoming the most favored application for determining adverse impact in employment discrimination cases.”).

⁴² See, e.g., *Isabel v. City of Memphis*, 404 F.3d 404, 411–13 (6th Cir. 2005); *Waisome v. Port Auth. of N.Y. & N.J.*, 948 F.2d 1370, 1375–80 (2d Cir. 1991); *Peightal v. Metro. Dade County*, 940 F.2d 1394, 1405–06 (11th Cir. 1991).

Sixth Circuit Court of Appeals—give little, if any, deference to the four-fifths rule.⁴³

As the courts continue to construe the four-fifths rule as having little validity, this federal “guideline” can no longer safely be used as a guideline for employers in administering promotional examinations. Neither the Supreme Court nor Congress has addressed the current relevance of the four-fifths rule. Because the purpose behind the four-fifths rule was to provide guidance to employers in administering non-discriminatory tests, and because the four-fifths rule is no longer a valid guide, the Supreme Court or Congress should restore the purpose of the four-fifths rule and either afford it greater deference or amend the Uniform Guidelines with a more appropriate and accurate statistical guideline. Allowing the four-fifths rule to remain in this current ambiguous state only frustrates the purpose of Title VII and adversely impacts both employers and employees.

1. Calculations Under the EEOC’s Four-Fifths Rule

To determine whether the four-fifths rule has been met, first:

calculate the rate of selection for each group by dividing the number of persons selected from a group by the number of applicants from that group; [then], observe which group has the highest selection rate; [and finally], calculate the impact ratios by comparing the selection rates for the other individual groups to that of the group having the highest selection rate.⁴⁴

A ratio of less than four-fifths (80%) may indicate an adverse impact under the EEOC Guidelines.⁴⁵

For example, if an employer has 120 applicants, 80 white and 40 African Americans, from which 60 are selected, 48 whites and 12 African Americans, it has a selection rate for white applicants of 48/80, or 60%.

⁴³ See, e.g., *Isabel*, 404 F.3d at 411–13; *Clady v. County of L.A.*, 770 F.2d 1421, 1428–29 (9th Cir. 1985).

⁴⁴ Adoption of Questions and Answers To Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures, 44 Fed. Reg. 11996, 11998 (Mar. 2, 1979).

⁴⁵ See 29 C.F.R. § 1607.4(D) (2008).

The selection rate for African-American applicants is 12/40, or 30%. According to the EEOC's four-fifths rule, the employer's selection procedures have an adverse impact on African Americans because their selection rate is equal to only 50% of the selection rate of whites; thus, it fails to meet the four-fifths or 80% threshold.⁴⁶

D. Typical Promotional Examinations and the Use of Cutoff Scores

The purpose of promotional examinations is to provide a reliable measure of employee abilities by which employers can make informed decisions as to who the best candidate is for a promotion. For an agency such as a city police department, "Civil service policy favors promotions . . . on the basis of tested qualifications."⁴⁷ It is usually the "function of the civil service commission to fix the standards by which the qualifications of applicants for appointments to civil service positions may be tested."⁴⁸

Typical promotional examinations include physical components (especially with police and fire personnel) and a written test on specific knowledge of job content. Often, the employer will set a specific cutoff score, which must be achieved in order to continue through the promotional evaluation process. According to the Code of Federal Regulations: "Where cutoff scores are used, they should normally be set so as to be reasonable and consistent with normal expectations of acceptable proficiency within the work force."⁴⁹

Generally "[t]o validate a cutoff score, the inference must be drawn that the cutoff score measures minimal qualifications."⁵⁰ However, there appears to be a split in the circuit courts of appeals with regards to setting cutoff scores at "minimal qualifications." Some circuits require cutoff

⁴⁶ See 5 Emp. Coord. Employment Practices § 23.29; see also Michael T. Kirkpatrick, *Employment Testing: Trends and Tactics*, 10 EMP. RTS. & EMP. POL'Y J. 623, 628 (2006); Elaine W. Shoben, *Differential Pass-Fail Rates in Employment Testing: Statistical Proof Under Title VII*, 91 HARV. L. REV. 793, 805-06 (1978).

⁴⁷ 15A AM. JUR. 2D *Civil Service* § 58 (2000) (citing *McAdams v. Barbieri*, 123 A.2d 182 (Conn. 1956)).

⁴⁸ 67 C.J.S. *Officers* § 81 (2007) (citing *Cowen v. Reavy*, 28 N.E.2d 390 (N.Y. 1940); *Frishein v. Reavy*, 33 N.Y.S.2d 894 (App. Div.), *aff'd* 45 N.E.2d 921 (N.Y. 1942); *Vitale v. Conway*, 98 N.Y.S.2d 902 (Sup. Ct. 1950), *aff'd* 105 N.Y.S.2d 403 (App. Div. 1951)).

⁴⁹ 29 C.F.R. § 1607.5(H) (2008).

⁵⁰ *Isabel v. City of Memphis*, 404 F.3d 404, 413 (6th Cir. 2005).

scores to be set at minimal qualifications,⁵¹ while others do not.⁵² After *Isabel*, employers within the Sixth Circuit Court of Appeals must ensure that cutoff scores are set at “minimal qualifications.”⁵³

Some employers may, out of fear of a possible disparate impact claim, opt to eliminate a promotional examination altogether. However, public sector employers may, in addition to losing the benefits of holding a promotional examination, violate state law requiring promotional examinations. This is especially true for safety-sensitive positions, such as police officers or firefighters.

For example, in Ohio, whenever a vacancy occurs in a position above the rank of patrol officer or firefighter, the municipal or township civil service commission is required, within a specified period, to hold a competitive promotional examination.⁵⁴ Such promotional examinations must be designed to test the relative capacity of the persons examined to discharge the duties of the position sought.⁵⁵ Although the civil service commission should be the one to determine the substantive issues for a civil service promotional examination,⁵⁶ the examinations must be in writing, unless an examination is for a position requiring the operation of machines or equipment, in which case the examination may include a practical demonstration test of the operation of such machines and equipment.⁵⁷

III. RECENT DEVELOPMENTS

A. *Current Sixth Circuit Law and the Isabel v. City of Memphis Decision*

A recent Sixth Circuit case highlights the current problems associated with the four-fifths rule and the need for legislative or judicial clarity on its

⁵¹ See, e.g., *id.*; *Lanning v. Se. Pa. Transp. Auth. (SEPTA)*, 181 F.3d 478, 489 (3d Cir. 1999).

⁵² See, e.g., *Bew v. City of Chi.*, 252 F.3d 891, 895 (7th Cir. 2001); *Bryant v. City of Chi.*, 200 F.3d 1092, 1099–1100 (7th Cir. 2000); *Ass’n of Mexican-Am. Educators v. State of Cal.*, 231 F.3d 572, 590 (9th Cir. 2000); *Guardians Ass’n of the N.Y. City Police Dept., Inc. v. Civil Serv. Comm’n of City of N.Y.*, 630 F.2d 79, 105 (2d Cir. 1980).

⁵³ *Isabel*, 404 F.3d at 413.

⁵⁴ See OHIO REV. CODE ANN. §§ 124.44–45 (LexisNexis 2007).

⁵⁵ See *id.* § 124.23.

⁵⁶ *State ex rel. Fern v. City of Cincinnati*, 832 N.E.2d 106, 117 (Ohio Ct. App. 2005).

⁵⁷ See § 124.45; see also 15A AM. JUR. 2D *Civil Service*, *supra* note 47, § 58.

use. In *Isabel v. City of Memphis*,⁵⁸ the city's police department implemented a promotional examination for sergeants competing for the promotion to lieutenant in 2000. The promotional examination consisted of four parts: "(1) a written test, (2) a practical exercise test, (3) performance evaluations for the previous two years, and (4) seniority points."⁵⁹ The four components of the test "would account for twenty percent, fifty percent, twenty percent, and ten percent, respectively, of each candidate's [overall] score."⁶⁰ The candidates had to first pass the written test in order to proceed to the next three components of the test.⁶¹ The city hired an industrial organizational psychologist, Dr. Mark Jones, to develop the written test for the promotional examination.⁶² With the assistance of certain trained Memphis police officers, Dr. Jones and the officers "estimated the percentage of minimally qualified candidates who could answer the questions correctly in order to impose a cutoff score."⁶³ Based upon this cutoff score, the plaintiffs, four African-American sergeants, sued the city alleging discrimination in violation of Title VII.⁶⁴

It is important to note the history behind the city's use of the cutoff score in its promotional examination. Prior to administering the promotional examination, "[T]he [c]ity had entered into a Memorandum of Understanding with the police department, [whereby] it was agreed that in order to be eligible for a promotion to lieutenant, a candidate must achieve a passing score of seventy on a written job knowledge test."⁶⁵ In 1996, Dr. Jones did not use the cutoff score in administering the promotional examination, despite the Memorandum of Understanding.⁶⁶ As a result of failing to use the cutoff score, the police officers' union filed a grievance, alleging that Dr. Jones had violated the Memorandum of Understanding.⁶⁷

In administering the 2000 promotional examination, Dr. Jones complied with the Memorandum of Understanding and failed all

⁵⁸ 404 F.3d 404 (6th Cir. 2005).

⁵⁹ *Id.* at 408.

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² *Id.*

⁶³ *Id.*

⁶⁴ *Id.*

⁶⁵ *Id.*

⁶⁶ *Id.*

⁶⁷ *Id.*

candidates who scored below the agreed-upon cutoff score of seventy on the written examination.⁶⁸ However, by strictly adhering to the Memorandum of Understanding's cutoff score of seventy, not enough minority candidates, in proportion to non-minority candidates, passed the written test to satisfy the four-fifths rule.⁶⁹ Thus, to avoid the adverse impact, and to meet the four-fifths standard, Dr. Jones lowered the cutoff score from seventy to sixty-six.⁷⁰ Although the new cutoff score did not result in adverse impact under the EEOC's four-fifths rule, under alternative statistical analyses—including the *t*-test (which measures the difference in mean scores between minority and non-minority candidates) and *z*-score (which measures statistical success across groups)—adverse impact was shown.⁷¹ While the new cutoff score passed the four-fifths rule, the Sixth Circuit relied on the alternative statistical analyses presented and held that the promotional examination was discriminatory and in violation of Title VII.⁷² In fact, the court stated: “[N]otwithstanding a test’s compliance with the four-fifths rule, other analyses may still reveal an adverse impact. . . . [W]e are grateful for statistics beyond the four-fifths rule analysis because we prefer to look to the sum of statistical evidence to make a decision in these kinds of cases.”⁷³ *Isabel* apparently overruled the earlier Sixth Circuit opinion of *Black v. City of Akron*,⁷⁴ in which the court held that alternative statistical analyses would not be considered by the court if the four-fifths rule had been passed.⁷⁵

As a result of the *Isabel* decision, it is clear that employers in the Sixth Circuit can no longer rely on the EEOC's Guideline that having a selection rate that passes the four-fifths rule “will generally not be regarded by Federal enforcement agencies as evidence of adverse impact.”⁷⁶ In fact, the EEOC's Guideline, as far as the Sixth Circuit is concerned, should now

⁶⁸ *Id.*

⁶⁹ *Id.* at 409.

⁷⁰ *Id.* Dr. Jones also eliminated nine of the test questions which he deemed to be “faulty,” and gave all candidates credit for answering the questions correctly in order to maintain a one hundred-point scale. *Id.*

⁷¹ *Id.*

⁷² *Id.* at 414.

⁷³ *Id.* at 412–13.

⁷⁴ 831 F.2d 131 (6th Cir. 1987).

⁷⁵ *Id.* at 134.

⁷⁶ 29 C.F.R. § 1607.4(D) (2008).

read: “Whether an employer’s selection rate passes the four-fifths rule is irrelevant; the employer must pass, at a minimum, the statistical analyses of a *t*-test and *z*-score in order to avoid a claim of adverse impact under Title VII.”⁷⁷

In short, “employers in this [c]ircuit no longer have that assurance that their scrupulous adherence to the four-fifths rule will insulate them from an adverse impact suit.”⁷⁸ Furthermore, as the dissent in *Isabel* correctly forewarned:

The Majority’s opinion opens the door for plaintiffs to file a Title VII suit based on the newest statistical flavor of the month, but provides absolutely no guidance as which tests are to be used in assessing whether an employment practice results in an adverse impact, how many tests are to be applied, or how to determine their statistical significance.⁷⁹

The *Isabel* decision, which offered no clear guidance to employers after stripping the effect of the four-fifths rule, has now invited future litigation to refine what statistical tests and significance levels actually constitute an adverse impact. One can expect to see an increase in the use of statistical experts; after all, whether the *t*-test and *z*-score could sustain the conflicting evidence from say, a chi-square analysis, has yet to be determined. To be safe, and until the ambiguity created by the *Isabel* decision is clarified either legislatively or judicially, employers in the Sixth Circuit should make sure their promotional examinations satisfy—at a minimum—the four-fifths rule, the *t*-test, and the *z*-score.

B. Split in the Circuits Regarding Deferential Weight Given to the Four-Fifths Rule

In further support of the need for legislative or judicial clarity as to the relevance of the four-fifths rule, there appears to be a split in the circuits regarding the deferential weight given to the four-fifths rule. For example, the First Circuit held that the four-fifths rule was a useful benchmark in determining disparate impact in *Boston Police Superior Officers*

⁷⁷ See *Isabel*, 404 F.3d at 409.

⁷⁸ *Id.* at 418 (Batchelder, J., dissenting).

⁷⁹ *Id.*

Federation v. City of Boston.⁸⁰ The Second Circuit relied on the four-fifths rule in determining disparate impact in *Pietras v. Board of Fire Commissioners*.⁸¹ Similarly, in the Fifth Circuit, a “substantially disproportionate impact” cannot be shown if the selection rate exceeds the four-fifths rule.⁸² The Seventh Circuit, in *Allen v. City of Chicago*,⁸³ found disparate impact resulting from a police sergeant’s promotion examination by using the four-fifths rule.⁸⁴ The Eighth Circuit relied on the four-fifths rule in *Firefighters Institute for Racial Equality v. City of St. Louis, Missouri*,⁸⁵ finding adverse impact where the four-fifths rule was satisfied.⁸⁶ However, such reliance on the four-fifths rule was subsequently found to be inappropriate in *Eubanks v. Pickens-Bond Construction Company*⁸⁷ and *Mems v. City of St. Paul, Department of Fire and Safety Services* where the sample size of the group was too small to warrant a finding of disparate impact.⁸⁸ In addition, several federal district courts have followed suit and used the four-fifths rule in determining whether an employment practice had a disparate impact on a protected group of employees.⁸⁹

Conversely, several circuits now turn to alternative statistical analyses, other than the four-fifths rule, in determining whether a test results in

⁸⁰ 147 F.3d 13, 21 (1st Cir. 1998); *see also* *Cotter v. City of Boston*, 323 F.3d 160, 170 (1st Cir. 2003).

⁸¹ 180 F.3d 468 (2d Cir. 1999).

⁸² *Moore v. Sw. Bell Tel. Co.*, 19 Fair Empl. Prac. Cas. (BNA) 232, 234 (E.D. Tex. 1978), *aff’d*, 593 F.2d 607 (5th Cir. 1979).

⁸³ 351 F.3d 306 (7th Cir. 2003).

⁸⁴ *Id.* at 310 n.4; *see also* *Bew v. City of Chi.*, 252 F.3d 891, 893–94 (7th Cir. 2001); *United States v. City of Chi.*, 633 F.2d 1354, 1357 n.8 (7th Cir. 1981).

⁸⁵ 616 F.2d 350 (8th Cir. 1980).

⁸⁶ *Id.* at 356–57.

⁸⁷ 635 F.2d 1341, 1347–48, 1350 (8th Cir. 1980).

⁸⁸ 224 F.3d 735, 740 (8th Cir. 2000).

⁸⁹ *See* *Davidson v. Citizens Gas & Coke Util.*, 470 F. Supp. 2d 934, 940–41 (S.D. Ind. 2007); *Williams v. Sprint/United Mgmt. Co.*, 99 Fair Empl. Prac. Cas. (BNA) 1489, 1496 (D. Kan. 2006); *Bradley v. City of Lynn*, 433 F. Supp. 2d 157, 159 (D. Mass. 2006); *Thomas v. City of Evanston*, 610 F. Supp. 422, 427–28 (N.D. Ill. 1985); *Bigby v. City of Chi.*, 38 Fair Empl. Prac. Cas. (BNA) 844, 847 (N.D. Ill. 1984).

adverse impact.⁹⁰ For example, the First Circuit held in *Fudge v. City of Providence Fire Department*⁹¹ that “[w]here the size of the sample is small, however, the ‘four-fifths rule’ is not an accurate test of discriminatory impact.”⁹² In *Waisome v. Port Authority of New York and New Jersey*,⁹³ the Second Circuit used the standard deviation measure to statistically determine whether there was disparate impact.⁹⁴ In *Spence v. City of Philadelphia*,⁹⁵ the Third Circuit denied a disparate impact claim despite a showing of a four-fifths rule violation.⁹⁶ In *Anderson v. Westinghouse Savannah River Co.*,⁹⁷ the Fourth Circuit did not rely on—and made no reference to—the four-fifths rule in determining disparate impact.⁹⁸ As noted above, the Sixth Circuit has also looked to statistical analyses other than the four-fifths rule in determining disparate impact under Title VII.⁹⁹ The Eighth Circuit did not even look to the four-fifths rule in *Eubanks v. Pickens-Bond Construction Co.*,¹⁰⁰ and noted that statistics “come in infinite variety and, like any other kind of evidence they may be rebutted. In short, their usefulness depends on all of the surrounding facts and circumstances.”¹⁰¹ The Ninth Circuit held in *Clady v. Los Angeles County*¹⁰² that it was not legally bound to the EEOC’s four-fifths rule because such “regulations do not have the force of law”¹⁰³ and

⁹⁰ For a critical analysis of the four-fifths rule in disparate impact cases, see Shoben, *supra* note 45, at 805–06.

⁹¹ 766 F.2d 650, 659 n.10 (1st Cir. 1985).

⁹² *Id.* at 659 (citing *Shoben*, *supra* note 46).

⁹³ 948 F.2d 1370 (2d Cir. 1991).

⁹⁴ *Id.* at 1376.

⁹⁵ 147 F. App’x 289, 292 (3d Cir. 2005); see also *El v. Se. Pa. Transp. Auth. (SEPTA)*, 479 F.3d 232, 244 (3d Cir. 2007) (stating that the Equal Employment Opportunity Commission’s Guidelines are not to be afforded great deference).

⁹⁶ See *id.* at 292.

⁹⁷ 406 F.3d 248 (4th Cir. 2005).

⁹⁸ See *id.*

⁹⁹ *Isabel v. City of Memphis*, 404 F.3d 404, 409 (6th Cir. 2005).

¹⁰⁰ 635 F.2d 1341 (8th Cir. 1980).

¹⁰¹ *Id.* at 1347 (citing *Int’l Bhd. of Teamsters v. United States*, 431 U.S. 324, 340 (1977)).

¹⁰² 770 F.2d 1421 (9th Cir. 1985).

¹⁰³ *Id.* at 1428 (citing *Gen. Elec. Co. v. Gilbert*, 429 U.S. 125, 141–142 (1976); *Aquilera v. Cook County Police and Corr. Merit Bd.*, 760 F.2d 844, 847 (7th Cir. 1985)).

that the four-fifths rule “has been sharply criticized by courts and commentators.”¹⁰⁴ Finally, the Eleventh Circuit¹⁰⁵ and the D.C. Circuit¹⁰⁶ both found alternative statistical analyses other than the four-fifths rule to support a disparate impact claim.

As stated earlier, the Supreme Court has not heard a Title VII disparate impact case since its ruling in *Wards Cove*.¹⁰⁷ Even if Congress will not step in and clarify the relevance of the EEOC’s four-fifths rule, the split in the circuits justifies the Supreme Court stepping in and granting certiorari to resolve the split in the circuits.

C. *Showing Statistical Significance with Alternative Statistical Methods*

As discussed above, courts apparently are turning to alternative statistical methods, other than the four-fifths rule, in determining whether or not an employment practice, such as a promotional examination, results in a disparate impact on a protected group of employees. There appears to be no established limit as to what statistical methods may be proffered to prove disparate impact. This section will outline some of the more prevalent methods being used today. The comparative hypothetical under Section IV will then show how the statistics are deemed “significant” (i.e., support a disparate impact claim) depending on which statistical method is used.

1. *The Standard Deviation Analysis*

Loosely defined, standard deviation analysis is a statistical method of determining whether the characteristic composition of an employer’s workforce is due to chance or some other factor, such as race or gender. The Supreme Court first relied on the use of the standard deviation analysis in a 1977 jury discrimination case.¹⁰⁸ That same year, the Supreme Court

¹⁰⁴ *Clady*, 770 F.2d at 1428; *see also* *Paige v. California*, 233 F. App’x. 646 (9th Cir. 2007) (applying a standard deviation statistical measure, rather than the four-fifths rule, in determining disparate impact).

¹⁰⁵ *See* *Peightal v. Metro. Dade County*, 940 F.2d 1394, 1405–06 (11th Cir. 1991).

¹⁰⁶ *See* *Reynolds v. Sheet Metal Workers Local 102*, 498 F. Supp. 952, 966–69 (D.D.C. 1980), *aff’d*, 702 F.2d 221 (D.C. Cir. 1981); *see also* *Wilmore v. City of Wilmington*, 533 F. Supp. 844, 851–54 (D. Del. 1982); *Rich v. Martin-Marietta*, 467 F. Supp. 587, 611–12 (D. Col. 1979).

¹⁰⁷ *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989).

¹⁰⁸ *See* *Castaneda v. Partida* 430 U.S. 482, 495–96 (1977).

relied on the standard deviation analysis in deciding an employment discrimination case.¹⁰⁹ The standard deviation analysis is now “one of the most widely accepted statistical methods of determining disparate impact”¹¹⁰

Specifically, “[a] standard deviation is the square root of the product of the total number in the sample, multiplied by the probability of selecting a member of the protected class, [multiplied] by the probability of selecting a majority member.”¹¹¹ Many courts, including the United States Supreme Court in *Castaneda v. Partida*,¹¹² have held that a disparity of two or three standard deviations may be statistically significant.¹¹³ “However, some courts have found two or three standard deviations inadequate to support a finding of discrimination under the particular circumstances presented.”¹¹⁴

The following hypothetical illustrates how to calculate the standard deviation:

If the percentage of protected group members in the relevant labor market (as derived from census or other data) is 30, the number which would be expected to occur in a work force of 1,000 is 300 (1,000 x 0.3). Thus, the standard deviation, calculated as the square root of the product of the total number of the sample (1,000) multiplied by the probability of selecting a protected group member (0.3) times the probability of selecting a majority member (0.7) is approximately 14.5. If the observed number of protected group members in the employer’s

¹⁰⁹ See *Hazelwood Sch. Dist. v. United States*, 433 U.S. 299, 308–12 (1977).

¹¹⁰ ROBERT BELTON ET AL., *EMPLOYMENT DISCRIMINATION LAW: CASES AND MATERIALS ON EQUALITY IN THE WORKPLACE* 226 (7th ed. 2004).

¹¹¹ 8 Emp. Coord. Employment Practices § 107:127 (2008) (footnote omitted).

¹¹² 430 U.S. at 496 n.17.

¹¹³ See RAMONA L. PAETZOLD & STEVEN L. WILLBORN, *THE STATISTICS OF DISCRIMINATION: USING STATISTICAL EVIDENCE IN DISCRIMINATION CASES* § 4.13 (Thomson West 2006) (citations omitted).

¹¹⁴ 45C AM. JUR. 2D *Job Discrimination* § 2431 (2002) (citing *Gay v. Waiters’ and Dairy Lunchmen’s Union, Local No. 30*, 694 F.2d 531, 551–52 (9th Cir. 1982); *Kilgo v. Bowman Transp., Inc.*, 789 F.2d 859, 872–73 (11th Cir. 1986) (district court’s finding of adverse impact among the various labor pools examined, ranging in disparity between 1.92 and 5.3 standard deviations, was not reversible error)).

work force is 285, the difference between the expected (300) and observed number of protected group members is approximately one standard deviation. If the protected group members number 264, the difference between their expected and observed numbers measures approximately 2.5 standard deviations.¹¹⁵

2. *The t-Test Analysis*

The independent *t*-test analysis determines whether two sample means differ significantly from each other (e.g., comparing the scores of African-American applicants with the scores of white applicants).¹¹⁶ Basically, the *t*-test is calculated by first determining the numerator (sample mean score from one group of applicants minus the sample mean score from the other group of applicants). Then, determine the denominator by taking the square root of the following two equations: (1) the standard deviation squared for the first group, divided by population of that group; plus, (2) the standard deviation squared for the second group, divided by the population of that group. The result is your *t*-test score which is then compared to a significance level chart¹¹⁷ to determine whether the disparity between the two groups is statistically significant. The chart compares levels of significance (e.g., .05 level) with the degrees of freedom for the group as a whole ($df = n1 + n2 - 2$).¹¹⁸

3. *Other Methods for Determining Statistical Significance*

Another method used for determining statistical significance in disparate impact cases is multiple regression analysis. Multiple regression analysis is a statistical method “designed to estimate the effects of several independent variables on a single dependant variable”¹¹⁹ (e.g., race, gender, etc.). The level of significance is measured by the “multiple correlation coefficient (*R*) [which] indicates the relationship between the criterion and

¹¹⁵ 8 Emp. Coord. Employment Practices, *supra* note 111, § 107:127.

¹¹⁶ See JERRY R. THOMAS & JACK K. NELSON, RESEARCH METHODS IN PHYSICAL ACTIVITY 137 (4th ed. 2001).

¹¹⁷ See Appendix *infra* for chart.

¹¹⁸ *Id.*

¹¹⁹ BELTON, *supra* note 110, at 228. The Supreme Court endorsed the use of multiple regression analysis in *Bazemore v. Friday*, 478 U.S. 385, 400 (1986).

a weighted sum of the predictor variables.”¹²⁰ The use of more than one predictor variable will generally increase the accuracy of the prediction.¹²¹

Another method is the chi-square analysis which measures deviations from expected behavior and determines if such deviations are a result of chance or some other factor, such as race or gender.¹²²

The chi-square value for any two series is determined from a standard statistical table. Where a discrepancy becomes larger than that number, it means that the differences have not occurred as a result of chance alone. However, if the chi-square value becomes smaller than that number, the change is probably a result of chance variations.¹²³

A *z*-score is another method used to show statistical significance and it basically “measures statistical success across groups.”¹²⁴ In addition to the *t*-test, the Sixth Circuit looked to the *z*-score in determining whether the promotional examination in the *Isabel* case had a disparate impact.¹²⁵ Specifically, the *z*-score is defined as “[t]he basic standard score that converts raw scores to units of standard deviation in which the mean is zero and a standard deviation is 1.0.”¹²⁶ The *z*-score formula is $z = (X - M)/s$.¹²⁷

IV. ANALYSIS

Until Congress clarifies the relevance of the EEOC’s four-fifths rule, or the Supreme Court grants certiorari to resolve the split in the circuits regarding the deferential weight afforded the four-fifths rule, employers and employees will continue to be harmed—and the purposes of Title VII will continue to be undermined. For example, suppose that in State A, an employer’s test which passes the four-fifths rule is found to be valid, whereas, that employer’s same test which happens to be administered at its sister plant in State B, while still passing the four-fifths rule, is found to not

¹²⁰ THOMAS & NELSON, *supra* note 116, at 127.

¹²¹ *Id.*

¹²² *See* NAACP v. City of Mansfield, 866 F.2d 162, 167 (6th Cir. 1989).

¹²³ *Id.*

¹²⁴ *Isabel* v. City of Memphis, 404 F.3d 404, 409 (6th Cir. 2005).

¹²⁵ *Id.*

¹²⁶ THOMAS & NELSON, *supra* note 116, at 191.

¹²⁷ *Id.* at 192.

defeat an adverse impact claim. The employer is now in a position to have to create two different methods for administering, and testing, its promotional examinations. These cost implications could prove to be significant.

Similarly, suppose that a *t*-test is a better statistical indicator of disparate impact than the four-fifths rule.¹²⁸ Assuming that to be true, the employees in the hypothetical State A would more likely be subjected to discrimination than would the employees in State B. Surely, in enacting Title VII, Congress did not intend the protections from discrimination to vary from state to state based on judicial interpretations of the four-fifths rule. To avoid this result, Congress or the Supreme Court must step in and clarify the current ambiguity surrounding the use of the four-fifths rule.

As the court stated in *Clady v. County of Los Angeles*:¹²⁹

There is no consensus on a threshold mathematical showing of variance to constitute substantial disproportionate impact. Some courts have looked to *Castaneda v. Partida*, which found adverse impact where the selection rate for the protected group was “greater than two or three standard deviations” from the selection rate of their counterparts.¹³⁰

The Supreme Court should grant certiorari, per the current split in the circuits, to hear a disparate impact claim under Title VII and use that opportunity to declare a standard statistical test to provide more uniformity and predictability in administering promotional examinations.

In *Watson v. Fort Worth Bank and Trust*,¹³¹ Justice O’Connor recognized that the Supreme Court has looked to two or three standard deviations to be used to show prima facie evidence of discrimination,¹³²

¹²⁸ See Shoben, *supra* note 46, at 805–06 (“The four-fifths rule is an ill-conceived resolution of the problem of assessing the substantiality of pass or acceptance rate differences. . . . These flaws in the four-fifths rule can be eliminated by replacing it with a test of the statistical significance of differences in pass rate proportions.”).

¹²⁹ 770 F.2d 1421 (9th Cir. 1985).

¹³⁰ *Id.* at 1428 (citing *Castaneda v. Partida*, 430 U.S. 482, 496–97 n.17 (1977)).

¹³¹ 487 U.S. 977 (1988).

¹³² *Id.* at 995–96 n.3. As defined in King:

(continued)

but that the Supreme Court has never instructed lower courts to do the same.¹³³ “We have emphasized the useful role that statistical methods can have in Title VII cases, but we have not suggested that any particular number of ‘standard deviations’ can determine whether a plaintiff has made out a prima facie case in the complex area of employment discrimination.”¹³⁴

As stated earlier, the Supreme Court has not heard a disparate impact claim since the passage of the Civil Rights Act of 1991. Such a case is bound to come to the Court soon, and the Supreme Court should take that opportunity to do what Justice O’Connor refused to do in *Watson*: instruct lower courts on the proper statistical proof to be used in disparate impact cases. If the Supreme Court refuses to provide this necessary guidance, Congress or the EEOC should revise the Guidelines to provide for a more definite rule than the four-fifths rule.

The comparative hypothetical below will best illustrate the problems resulting from the ambiguous status of the four-fifths rule and the use of alternative statistical analyses, and will highlight the need for legislative or judicial clarification on the four-fifths rule and statistical methods used to determine disparate impact.

A. A Comparative Hypothetical Using Alternative Statistical Analyses

As Justice Stewart stated in *Teamsters*,¹³⁵ “statistics . . . come in infinite variety.”¹³⁶ Courts within the Sixth Circuit Court of Appeals are now likely to see the variety of statistics as a result of *Isabel*. The comparative hypothetical below will illustrate the inconsistency in results

The “standard deviation” is a unit of measurement that allows statisticians to measure all types of disparities in common terms. Technically, a “standard deviation” is defined as “a measure of spread, dispersion, or variability of a group of numbers equal to the square root of the *variance* of that group of numbers.”

King, *supra* note 13, at 275 (citing D. BALDUS & J. COLE, STATISTICAL PROOF OF DISCRIMINATION 359 (1980)).

¹³³ King, *supra* note 13, at 276.

¹³⁴ *Watson*, 487 U.S. at 996 n.3.

¹³⁵ *Int’l Bhd. of Teamsters v. United States*, 431 U.S. 324 (1977).

¹³⁶ *Id.* at 340.

created by using different methods of statistical analysis. Keep in mind that “statistics often tell much, and [c]ourts listen.”¹³⁷

Assume the Human Resources Manager for the city of XYZ has been asked to implement a promotional examination for the lieutenant position in the city’s police department. The goal is to better evaluate the applicants, and, therefore, promote the most qualified officer—in a nondiscriminatory way. The city contracts with Dr. Smith, an Industrial/Organizational Psychologist, to develop an effective and valid nondiscriminatory promotional examination. Having referred to the federal government’s guidelines on discrimination, the Human Resources Manager cautions Dr. Smith that the promotional examination must withstand the EEOC’s four-fifths rule.¹³⁸

The promotional examination is developed and consists of the following four parts: (1) a written test, (2) a practical exercise test, (3) performance evaluations for the previous two years, and (4) seniority points. The four components are to account for twenty percent, fifty percent, twenty percent, and ten percent, respectively, of each applicant’s total score. Those who pass the written test will proceed to the next three components; those who do not pass will be disqualified from continuing the promotional process. The cutoff score for passing the written test component of the promotional examination is set at seventy points. The promotions will be offered to those applicants ranking at the top in total scores.

There were a total of 120 applicants for the promotion; 80 applicants were white and 40 were African American. A total of 60 applicants, 48 white and 12 African American, passed the written test. The average score (population mean) for the white applicants was 69.775 and 70.975 for the African-American applicants. One African-American applicant failed to pass the written test and has now sued the city claiming racial discrimination. What result?

1. Four-Fifths Rule Analysis

Under the four-fifths rule analysis, the plaintiff will be able to show that the city’s promotional examination was discriminatory. Because forty-eight of the eighty white applicants passed the written part of the test,

¹³⁷ Alabama v. United States, 304 F.2d 583, 586 (5th Cir. 1962).

¹³⁸ 29 C.F.R. § 1607.4(D) (2008).

the selection rate for white applicants was 60% (48/80). The selection rate for African-American applicants was 30% (12/40). The selection rate for the African-American applicants was equal to 50% of the selection rate for white applicants (30%/60%). The 50% difference fails to meet the four-fifths (80%) threshold; thus, the city's promotional examination had a disparate impact on African-American applicants with the cutoff score set at seventy.

2. *The t-Test Analysis*

The following baseline data is derived from the same data as was used to calculate the four-fifths rule analysis. The white applicants had a population (N) of 80; a population mean (m), or total average score, of 69.775; a median of 70; a mode of 61; a standard deviation (s) of 13.15; and, a sample mean (M), or average score of the passing applicants, of 80.79. The African-American applicants had a population (N) of 40; a population mean (m), or total average score, of 70.975; a median of 67; a mode of 69; a standard deviation (s) of 17.45; and, a sample mean (M), or average score of the passing applicants, of 94.58.

The first step in calculating the t -test is determining the degrees of freedom. "Degrees of freedom are based on the number of participants with a correction for bias: $df = n - 1$."¹³⁹ When analyzing two independent variables, such as with this comparative hypothetical, the equation for calculating degrees of freedom is $df = n_1 + n_2 - 2$.¹⁴⁰ Therefore, the degrees of freedom equals 118 ($80 + 40 - 2 = 118$).

Comparing the degrees of freedom of 118 to a .05 level of significance for the one-tailed test, the Critical Values of t Chart¹⁴¹ shows that in order for the t -test to be statistically significant (i.e., to show disparate impact), the value of t must be greater than or equal to 1.296.¹⁴² To calculate the t -test, first determine the numerator by subtracting the population mean (M) score for the white applicants from the population mean score for the African-American applicants. This results in a numerator of 1.2 ($70.975 - 69.775 = 1.2$). To calculate the denominator, take the square root of the standard deviation squared for the African-American applicants ($s =$

¹³⁹ THOMAS & NELSON, *supra* note 116, at 137.

¹⁴⁰ *Id.*

¹⁴¹ See Appendix *infra*.

¹⁴² THOMAS & NELSON, *supra* note 116, at 401.

17.45), divided by the African-American population ($n = 40$), plus the standard deviation squared for the white applicants ($s = 13.15$) divided by the white population ($n = 80$). The result for the denominator is 3.13 $((17.45)(17.45) = 304.5/40 = 7.6 + (13.15)(13.15) = 172.92/80 = 2.16$; square root of $7.6 + 2.16$ (9.77) equals 3.13). Dividing the numerator (1.2) by the denominator (3.13) results in a t -score of 0.38. Because 0.38 is not greater than the significance level of 1.296 for 118 degrees of freedom, there is no statistical significance and, therefore, no evidence of disparate impact on African-American applicants taking the promotional examination. Thus, whereas the four-fifths rule analysis showed disparate impact, the t -test showed there was no disparate impact.

3. Standard Deviation

The result under the standard deviation analysis becomes even more ambiguous. Again, the standard deviation is the “square root of the product of the total number in the sample, times the probability of selecting a member of the protected class, times the probability of selecting a majority member.”¹⁴³

Using the hypothetical numbers above, it would be expected that white applicants have 66.67% chance of passing the cutoff score (80 white applicants/120 total applicants = 66.67%); and African-American applicants would be expected to have a 33.34% chance of passing the cutoff score (40 African-American applicants/120 total applicants = 33.34%). Using the total number of passing applicants as the sample (60 applicants), the standard deviation equates to 3.65 (the square root of 60 applicants \times 33.34% \times 66.67% = 3.65).

Because the chance of African-American applicants passing the cutoff score is 33.34%, and a total of sixty applicants passed the cutoff score, it would be expected that twenty African-American applicants would pass the cutoff score (60 total applicants \times 33.34% = 20). However, as indicated above, in this hypothetical only twelve African-American applicants passed the cutoff score. Therefore, the difference between the expected number of passing African-American applicants (twenty) and the actual number of passing African-American applicants (twelve) is greater

¹⁴³ 45C AM. JUR. 2D *Job Discrimination*, *supra* note 114, § 2431 (citing *Garrett v. R.J. Reynolds Indus. Inc.*, 81 F.R.D. 25 (M.D.N.C. 1978)).

than two standard deviations (20 expected – 12 actual = 8/3.65 standard deviation = a difference of 2.19 standard deviations).

Although the Supreme Court has acknowledged “two or three standard deviations” as being statistically significant,¹⁴⁴ such guidance does not help determine whether the hypothetical promotional examination was discriminatory, since its result of 2.19 standard deviations falls between two and three standard deviations.

In other words, if the court hearing the hypothetical case above held two standard deviations to be significant, the plaintiff would be able to show disparate impact and would likely win. However, if the court held three standard deviations was necessary to show statistical significance, the plaintiff would fail to show disparate impact and would likely lose.

B. Problems Highlighted by the Comparative Hypothetical

The comparative hypothetical highlights some of the major problems that employers, employees, and courts under the Sixth Circuit Court of Appeals must face, given the ambiguous status of the four-fifths rule. Recall that in the *Isabel* decision, the Sixth Circuit Court of Appeals found evidence of disparate impact based on the results of the *t*-test analysis, even though there was no evidence of disparate impact under the four-fifths rule analysis.¹⁴⁵ However, in the comparative hypothetical, there was no evidence of disparate impact under the *t*-test analysis; yet there was evidence of disparate impact under the four-fifths rule analysis. How should a district court within the Sixth Circuit’s jurisdiction rule on the facts given in the comparative hypothetical? How would the Sixth Circuit rule if it were faced with such facts?

The standard deviation analysis does not provide much more guidance in resolving the difference posed by the four-fifths rule and *t*-test results. Because the difference between the expected value and the actual value was 2.19 standard deviations, and because the Supreme Court has only stated that “two *or* three standard deviations” may be statistically

¹⁴⁴ *Hazelwood Sch. Dist. v. United States*, 433 U.S. 299, 309 n.14 (1977). Because the Court was faced with evidence of discrimination by way of six standard deviations, it was not necessary for the Court to set a bright-line rule of whether two *or* three standard deviations qualified as being statistically significant. *Id.*

¹⁴⁵ *Isabel v. City of Memphis*, 404 F.3d 404, 409 (6th Cir. 2005).

significant,¹⁴⁶ again, it is unclear how the Sixth Circuit, and its lower courts, should or would rule.

Although no statistical evidence under the standard deviation analysis was presented to the Sixth Circuit in the *Isabel* decision,¹⁴⁷ the court did invite the use of statistical evidence beyond the use of the four-fifths rule.¹⁴⁸ With this being said, it is interesting to note that had the Sixth Circuit relied on the standard deviation analysis, rather than the *t*-test analysis, it would have found no evidence of disparate impact, consistent with the results of the four-fifths rule.

The facts and numbers of the comparative hypothetical are similar to those of the *Isabel* case. In the *Isabel* case, there were a total of 120 applicants for the promotional examination; fifty-seven were white and sixty-three were African-American.¹⁴⁹ Therefore, the white applicants had a 48% expected chance of passing the cutoff score (57 total white applicants/120 total applicants = 48%); and, the African-American applicants had approximately a 52% chance of passing (63 total African-American applicants/120 total applicants = 52%). The standard deviation for this sample is 5.0 (square root of 98 total passing applicants x 52% expected chance for African-American applicants passing x 48% expected chance for white applicants passing = 5.0).

Because ninety-eight applicants passed the cutoff score, the expected number of African-American applicants to pass would have been fifty-one (98 total passing applicants x 52% expected chance of passing = 51). However, only forty-seven African-American applicants passed the cutoff score. Because the standard deviation for this sample is 5.0, the difference between the expected value and the actual value is less than one standard deviation (51 expected to pass – 47 that actually passed = 4). Therefore, even under the Supreme Court's significance standard of "two or three standard deviations,"¹⁵⁰ this result is not statistically significant; thus, it

¹⁴⁶ *Hazelwood Sch. Dist.*, 433 U.S. at 309 n.14 (1977) (emphasis added).

¹⁴⁷ *Isabel*, 404 F.3d at 410 (6th Cir. 2005).

¹⁴⁸ *Id.* at 412–13 (“[W]e are grateful for statistics beyond the four-fifths rule analysis because we prefer to look to the sum of the statistical evidence to make a decision in these kinds of cases.”).

¹⁴⁹ *Id.* at 408.

¹⁵⁰ *Hazelwood Sch. Dist.*, 433 U.S. at 309 n.14 (quoting *Castaneda v. Partida*, 430 U.S. 482, 497 n.17 (1977)).

supports the opposite of the *Isabel* court's holding—there was no disparate impact and the promotional examination was not discriminatory. How would the Sixth Circuit have ruled in *Isabel* had it also been presented with statistics under the standard deviation analysis? As the dissenting opinion forewarned, the result may have depended on whatever was the “newest statistical flavor of [that] month.”¹⁵¹

Given the uncertainty in the courts, employers will likewise be uncertain in determining whether they have developed a non-discriminatory promotional examination; and employees will also be uncertain as to whether they should expend the time and money necessary to bring a disparate impact lawsuit.

C. Possible Solutions to the Problems with the Four-Fifths Rule

Until the current state of confusion is settled, the EEOC should, at the very least, amend the four-fifths rule to eliminate the provision stating that passing the four-fifths rule will generally not be regarded as evidence of adverse impact.¹⁵² This is simply not the case; and providing such a guideline is misleading to those who seek to develop a non-discriminatory promotional examination. Making such a revision begs the question: what statistical method should replace the four-fifths rule analysis?

Although the Supreme Court has not yet instructed lower courts to apply a test of two or three standard deviations,¹⁵³ perhaps now is the time. The standard deviation analysis has been recognized by the Supreme Court since the *Castaneda v. Partida*¹⁵⁴ and *Hazelwood School District v. United States*¹⁵⁵ decisions, and the “two or three standard deviation” test has been utilized by several lower courts.¹⁵⁶ Furthermore, similar to the *t*-test,¹⁵⁷ two

¹⁵¹ *Isabel*, 404 F.3d at 418 (Batchelder, J., dissenting).

¹⁵² 29 C.F.R. § 1607.4(D) (2008).

¹⁵³ *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977, 995 n.3 (1988) (discussing that a suggested number of standard deviations has not been established).

¹⁵⁴ 430 U.S. 482 (1977).

¹⁵⁵ 433 U.S. 299 (1977).

¹⁵⁶ *King*, *supra* note 13, at 277 (quoting *Peightal v. Metro. Dade County*, 940 F.2d 1394, 1406 (11th Cir. 1991) (citations omitted) (“The ‘general rule’ is that the disparity must be ‘greater than two or three standard deviations’ before it can be inferred that the employer has engaged in illegal discrimination under Title VII.”)); *see also* *Smith v. Xerox Corp.*, 196 F.3d 358, 365–66 (2d Cir. 1989) (finding that a disparity of two or three standard deviations raises an inference of discrimination); *Palmer v. Shultz*, 815 F.2d 84, 98–99

(continued)

standard deviations are roughly equivalent to a .05 confidence level and three standard deviations are roughly equivalent to a .01 confidence level—levels which have both been “cited with approval by courts as a proper method of measuring statistical significance.”¹⁵⁸ Given this, the Supreme Court should grant certiorari for an upcoming disparate impact case and finally declare the two or three standard deviation test to be the uniform statistical method for determining disparate impact.

If the Supreme Court refuses to announce such a definitive rule, Congress or the EEOC could also decide to replace the four-fifths rule with a test of two or three standard deviations. Whether the two or three standard deviation test is legislatively or judicially imposed, the end result will still be met—providing clarity and uniformity for employers, employees, and courts to rely on in determining whether a promotional examination has a disparate impact.

Finally, one last solution would be for the Supreme Court, Congress, or the EEOC to declare the four-fifths rule to be the mandatory statistical test used for determining disparate impact. Although the four-fifths rule may not be the most accurate measure of disparate impact, establishing it as the mandatory measure of disparate impact would provide the necessary clarity and uniformity. In short, whatever remedial provisions are employed, it is clear that the current ambiguous state of the four-fifths rule must not remain.

D. How Employers in the Sixth Circuit Can Survive the Ambiguity

Given the current ambiguous state of the four-fifths rule after the *Isabel* decision,¹⁵⁹ employers in the Sixth Circuit cannot haphazardly develop a promotional examination. Based on the cases and discussion

(D.C. Cir. 1987) (finding that a disparity of higher than three standard deviations raises an inference of discrimination); *cf. Gay v. Waiters' & Dairy Lunchmen's Union, Local No. 30*, 694 F.2d 531 (9th Cir. 1982) (quoting *EEOC v. AM. Nat'l Bank*, 652 F.2d 1176, 1192 (4th Cir. 1981) (“[C]ourts should be ‘extremely cautious’ of drawing any inferences from standard deviations in the range of 1 to 3.”)).

¹⁵⁷ See THOMAS & NELSON, *supra* note 116, at 142.

¹⁵⁸ 45C AM. JUR. 2D *Job Discrimination*, *supra* note 114, § 2431 (citing *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975); *Reynolds v. Sheet Metal Workers Local 102*, 498 F. Supp. 952 (D.D.C. 1980); *Vuyanich v. Republic Nat'l Bank of Dallas*, 505 F. Supp. 224 (N.D. Tex. 1980)).

¹⁵⁹ *Isabel v. City of Memphis*, 404 F.3d 404 (6th Cir. 2005).

above, employers should, at the very least, take the following precautions when developing and implementing their promotional examinations.

First, the employer should ensure that the overall promotional examination is job related. Second, in setting a cutoff score, which applicants must pass in order to proceed through the promotional process, the employer should set the cutoff score at the minimal qualifications for the job. Although this step may not be necessary in other jurisdictions, it is essential in the Sixth Circuit after the *Isabel* decision.¹⁶⁰

After the promotional examination is determined to be job related, and the cutoff score is set at the position's minimum qualifications, the employer must run a statistical analysis of the results of those applicants passing the cutoff score (as well as those applicants passing the overall examination). Whereas the four-fifths rule used to be a reliable guideline, now the employer must ensure that not only did the test results pass the four-fifths rule, it must also pass muster under the *t*-test and probably under the *z*-score. To be safe, the employer may also want to ensure that the results satisfy the three, if not the two, standard deviation test. If the results of the employer's promotional examination show no statistical significance under the statistical tests above, the employer is probably safe to conclude that the promotional examination has no disparate impact.

Although it is highly recommended that the employer consult a statistical expert for determining whether its test has a disparate impact on a protected group of applicants, there are several sources available for employers to become more familiar with the statistical analysis portion of the promotional examination.¹⁶¹ If an employer follows the above recommendations in developing its promotional examinations, it should probably survive the ambiguity posed by the *Isabel* decision.¹⁶² However,

¹⁶⁰ See *id.* at 413.

¹⁶¹ See, e.g., AM. PSYCHOLOGICAL ASS'N, STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING (1999); SOC'Y FOR INDUS. & ORGANIZATIONAL SOC'Y, PRINCIPLES FOR THE VALIDATION AND USE OF PERSONNEL SELECTION PROCEDURES (4th ed. 2003); DAVID C. BALDUS & JAMES W.L. COLE, STATISTICAL PROOF OF DISCRIMINATION §§ 4.1–.3 (1980); NORMAN H. NIE ET AL., SPSS: STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES (2d ed. 1975); Shoben, *supra* note 46; Michael O. Finkelstein, *The Judicial Recognition of Multiple Regression Studies in Race and Sex Discrimination Cases*, 80 COLUM. L. REV. 737 (1980); Kingsley R. Browne, *Statistical Proof of Discrimination: Beyond "Damned Lies,"* 68 WASH. L. REV. 477 (1993).

¹⁶² *Isabel*, 404 F.3d 404.

until the Supreme Court, Congress, or the EEOC steps in to settle the confusion, there are no guarantees.

V. CONCLUSION

As discussed above, the result of the *Isabel* decision is that employers in the Sixth Circuit can no longer rely on the EEOC's guidance of applying the four-fifths rule to the selection rate in promotional examinations. At the very least, employers should prepare to test their promotional examinations for disparate impact using the *t*-test and *z*-score. To hedge on the safe side, the employers should also test their promotional examinations under alternative statistical analyses, such as the standard deviation analysis, which most likely will be introduced soon to the disparate impact litigation arsenal in the Sixth Circuit.

In addition, due to the split in the circuits and problems associated with the current ambiguity behind the four-fifths rule, the Supreme Court, Congress, or the EEOC should clarify the relevance of the four-fifths rule, and adopt a more accurate statistical analysis to be used. This will lead to uniformity and predictability; and, it will help maintain the purpose behind the enactment of Title VII.

APPENDIX¹⁶³
CRITICAL VALUES OF *T*

Level of Significance for One-Tailed Test

<i>df</i>	0.1	0.05	0.025	0.01	0.005	0.0005
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

¹⁶³ Appendix derived from Thomas & Nelson, *supra* note 116, at 401.