Washington and Lee Law Review

Volume 71 | Issue 4 Article 7

Fall 9-1-2014

The Use and Misuse of Econometric Evidence in Employment Discrimination Cases

Joni Hersch
Vanderbilt Law School

Blair Druhan Bullock
Vanderbilt Law School

The Use and Misuse of Econometric Evidence in Employment Discrimination Cases

Joni Hersch*
Blair Druhan Bullock**

Abstract

Statistical analyses play an important role in employment discrimination cases, as the Supreme Court has long recognized. Regression analysis can help a plaintiff establish a claim of discrimination under Title VII of the Civil Rights Act of 1964 by showing that, even when controlling for relevant characteristics, individuals of a certain class were treated differently than other employees or applicants. It can also help a defendant rebut such a claim by showing that differential treatment was due to characteristics other than being a member of a protected class. Yet, too often, opposing experts present invalid rebuttal evidence that the jury or judge overweighs. Opposing experts routinely criticize three aspects of the regression: the regression's explanatory variables, its sample size, and its statistical significance. Even though these factors affect the reliability of the regression results only in very limited circumstances, the judge or jury is often persuaded by them and find for the defendant. As a result, valid regression analyses do not perform the critical work that they should in employment discrimination cases. Our own statistical analyses of seventy-eight Title VII employment discrimination cases finds that regression analyses do not

* Professor of Law and Economics at Vanderbilt Law School, Co-Director of the Ph.D. Program in Law and Economics.

** Ph.D. Candidate, Program in Law and Economics at Vanderbilt Law School. J.D. Candidate, Vanderbilt Law School.

We thank Lisa Bressman, Caroline Cecot, Benjamin McMichael, Michael Selmi, Jennifer Bennett Shinall, Kevin Stack, and Michael Vandenbergh for their valuable comments and Danielle Drago and Jean Xiao for research assistance.

substantially increase the plaintiff's likelihood of prevailing at trial and that if the court recognizes any of these common critiques, the plaintiff is much less likely to prevail. The severe consequences of such critiques make it very important for the court and opposing experts to recognize when these critiques are without merit. We propose that courts adopt a peer-review system in which court-appointed economists, compensated by each party as a percentage of the total payment to econometric expert witnesses, review econometric evidence before the reports are submitted to the judge or jury.

Table of Contents

I. Introduction

"Do you have a low-pitched voice? Do you swear often? Have you ever done any hunting? Have you participated in wrestling? Have you participated in boxing? Have you played football on a team?"[1] These were questions asked during the hiring process for sales representatives at Sears, Roebuck & Co. in the 1980s.[2] While these questions may appear to be on their face discriminatory, this evidence was not enough for a class of female employees to establish gender discrimination in hiring in *E.E.O.C. v. Sears, Roebuck & Co*.[3] To bolster its case, the plaintiff introduced regression analyses that showed that, controlling for important factors including job applied for, age, education, job-type experience, product-line experience, and commission-product experience, females were statistically less likely to be hired as sales representatives at Sears.[4] However, this statistical evidence did not improve the plaintiff's case, as the defendant challenged the regression analysis because it did not control for certain factors deemed by Sears to be desirable for sales representatives, including factors based on the above questions and "physical appearance, assertiveness, the ability to communicate, friendliness, and economic motivation."[5] Though the court

---

1. Ruth Milkman, *Women's History and the* Sears *Case*, 12 FEMINIST STUD. 375, 382 (1986) (quoting Plaintiff's Pretrial Brief at 34, E.E.O.C. v. Sears, Roebuck & Co., 628 F. Supp. 1264 (N.D. Ill. 1986) (No. 7964373)).

2. *See id.* (noting these questions were components of an applicant's vigor score, which was used to make hiring decisions).

3. *See* E.E.O.C. v. Sears, Roebuck & Co., 628 F. Supp. 1264, 1318 (N.D. Ill. 1986) ("There is no credible evidence that a woman's 'vigor' score ever prevented her from being hired into commission sales at Sears. The court therefore finds that Sears' testing program did not discriminate against women . . . .").

4. *See id.* at 1296 (discussing a weighted logit regression analysis that used these six factors).

5. *See id.* at 1303 ("Other important factors not controlled for in EEOC's analysis are those characteristics which could be determined only from an interview, not from the written application. These include physical appearance,

acknowledged these qualities were difficult to quantify when relying on this argument,[6] the court did not require the defendant to prove that these qualities varied with gender or to establish statistically their importance in hiring. In part because of the reliance on this invalid critique, the plaintiffs in this case were left without recourse.[7]

Parties involved in discrimination cases have presented statistical analyses to bolster their cases for decades.[8] In fact, the Supreme Court recognized the important role of statistical analyses in discrimination cases more than thirty-five years ago in *International Brotherhood of Teamsters v. United States*.[9] While statistical analyses and, in particular, regression analyses still maintain an important role in discrimination cases, that role continues to be diminished by rebuttal evidence presented by the opposing party.[10] Too often, this rebuttal evidence presents

---

assertiveness, the ability to communicate, friendliness, and economic motivation.").

6. *See id.* at 1303 n.34 ("The court recognizes that these factors are not easily quantified for purposes of a statistical analysis, and that data relating to these factors was generally not available to EEOC from the application forms it chose to rely upon.").

7. *See id.* at 1353 ("Accordingly, based on the above findings of fact and conclusions of law, it is hereby adjudged and ordered that judgment is entered against plaintiff and in favor of defendant on all claims at issue in the trial of this case, and plaintiff's claim for relief is hereby denied.").

8. *See, e.g.*, City of Richmond v. J.A. Croson Co., 488 U.S. 469, 509 (1989) ("Moreover, evidence of a pattern of individual discriminatory acts can, if supported by appropriate statistical proof, lend support to a local government's determination that broader remedial relief is justified."); Bazemore v. Friday, 478 U.S. 385, 387 (1986) (per curiam) (finding that the court of appeals erred by disregarding petitioners' statistical analyses even though the analyses reflected salary disparities in place before Title VII applied to the defendant); Furnco Constr. Corp. v. Waters, 438 U.S. 567, 580 (1978) (ruling that on remand the court must consider statistical evidence showing the employers' work force was racially balanced); United States v. City of New York, 637 F. Supp. 2d 77, 86 (E.D.N.Y. 2009) (explaining that statistical evidence of disparate impact may suffice to establish a prima facie case of discrimination).

9. *See* Int'l Bhd. of Teamsters v. United States, 431 U.S. 324, 339 (1977) ("In any event, our cases make it unmistakably clear that '(s)tatistical analyses have served and will continue to serve an important role' in cases in which the existence of discrimination is a disputed issue." (citation omitted)).

10. *See City of New York*, 637 F. Supp.2d at 85 (explaining ways a defendant can rebut statistical evidence presented by the plaintiff to establish a prima facie case of discrimination).

invalid critiques that the jury or judge overweighs. As a result, valid regression analyses are often incorrectly negated.

Proper regression analyses can serve an important role in employment discrimination cases. They can help a plaintiff establish a claim of discrimination under Title VII by showing that, even when controlling for relevant characteristics, individuals of a certain class were treated differently than other employees or applicants.[11] Alternatively, they can help a defendant rebut such a claim by showing that differential treatment was due to characteristics other than being a member of a protected class.[12] In addition, despite the Supreme Court's recognition in *Wal-Mart Stores, Inc. v. Dukes* that regression analyses may not always be appropriate,[13] regression analyses can still assist a class of plaintiffs trying to establish commonality. Such regression analyses establish that the entire class, as members of a protected class under Title VII, experienced the same form of discrimination. Unfortunately, due to incorrect challenges, often backed by expert witnesses, regression analyses do not always serve these important purposes.

All too often, once a party presents regression analyses to assist its case, the opposing party launches spurious critiques challenging the validity of the analyses.[14] Then, without critically evaluating those critiques, the judge either accepts the critiques

---

11. *See, e.g.*, Lavin-McEleney v. Marist Coll., 239 F.3d 476, 478 (2d Cir. 2001) (discussing a regression analysis that showed the plaintiff was paid less than male professors even after controlling for relevant factors such as experience, tenure status, and type of degree).

12. *See, e.g.*, Morgan v. United Parcel Serv. of Am., Inc., 143 F. Supp. 2d 1143, 1151 (E.D. Mo. 2000) (explaining how the defendant's expert argued a wage disparity was not based on race because, if the regression analysis controlled for all performance evaluations, then race was not a statistically significant factor).

13. *See* Wal-Mart Stores, Inc. v. Dukes, 131 S. Ct. 2541, 2555–56 (2011) (explaining that the regression analyses presented as evidence could not establish commonality because a regional disparity does not prove that each store within the region has the same disparity).

14. *See, e.g.*, Carpenter v. Boeing Co., 456 F.3d 1183, 1196 (10th Cir. 2006) (providing an example of an expert attacking a statistical analysis because variables were missing from the study, even though the expert did not demonstrate that the missing variables affected the statistical significance of the results).

as valid support for a motion or allows the critiques to enter the courtroom, where the critiques are overweighed by the jury. This often leads to an unbalanced discussion about everything potentially wrong with the analyses, instead of a discussion about their actual validity. For example, throughout the highly publicized litigation of *Dukes v. Wal-Mart Stores, Inc.*,[15] the plaintiffs' and defendant's experts debated whether the presented regression analysis established class commonality and provided evidence that Wal-Mart discriminated against female employees by paying them less.[16] As the Northern District of California noted in a full 25% of its class-certification motion, the defendant's expert claimed that the plaintiffs' regression analysis was invalid because it failed to separately analyze each division of each store and incorrectly analyzed the entire sample of employees within a region at once.[17]

This unbalanced discussion occurs frequently. All too often the opposing experts criticize three aspects of the regression: the regression's explanatory variables, its sample size, and its statistical significance, all of which affect the reliability of the regression results only in very limited circumstances.[18] By

---

15. 222 F.R.D. 137 (N.D. Cal. 2004), *rev'd*, 131 S. Ct. 2541 (2011).

16. *See id.* at 155 ("Plaintiffs present largely uncontested descriptive statistics which show that women working in Wal-Mart stores are paid less than men in every region, that pay disparities exist in most job categories, that the salary gap widens over time even for men and women hired into the same jobs . . . .").

17. *See id.* at 156 (explaining the defendant's contention that the statistical analysis at the regional level fails to account for significant differences in compensation practices among the individual stores). These arguments eventually led the Supreme Court of the United States to hold that the statistical evidence presented by the plaintiffs did not establish a company-wide policy of gender discrimination required for commonality and for class certification, establishing precedent limiting the use of regression analysis in class certification motions. *Wal-Mart Stores, Inc.*, 131 S. Ct. at 2555–56.

18. *See, e.g.*, Franklin v. Local 2 of the Sheet Metal Workers Int'l Assoc., 565 F.3d 508, 514 (8th Cir. 2009) (providing an example of criticism based on variables omitted from the regression that may alter the results); Coleman v. Exxon Chem. Corp., 162 F. Supp. 2d 593, 618 (S.D. Tex. 2001) (providing an example of criticism based on a sample size of forty individuals even though eight individuals were members of the relevant protected class); Boyd v. Interstate Brands Corps., 256 F.R.D. 340, 361 (E.D.N.Y. 2009) (providing an example of an expert challenging a plaintiff's statistical analysis because of statistical significance).

focusing on these presented econometric criticisms, the judge or jury is often persuaded that this evidence is not reliable, and as a result, strong and valid evidence of discrimination is disregarded, and the defendant prevails.[19] This Article analyzes the presentation of these critiques in Title VII employment discrimination cases and proposes ways for the court to avoid allowing an unbalanced discussion of potential econometric critiques to negate such valuable evidence. Our own statistical analyses of seventy-eight published employment discrimination cases finds that regression analyses do not increase substantially the plaintiff's likelihood of prevailing at trial and that, if the court recognizes any of these common critiques, the plaintiff is much less likely to prevail. The severe consequences of such critiques make it even more important for the court and for opposing experts to recognize when these critiques themselves are without merit.

This Article begins by discussing how regression analyses are presented in employment discrimination cases and by analyzing the court's recognition of the potential problems with the analyses. Part III discusses three of the most common, invalid econometric critiques found in employment discrimination cases: omitted variables, sample size deficiencies, and lack of statistical significance. Part III also establishes the rare circumstances when these critiques are actually valid. Part IV then presents a statistical analysis of published employment discrimination cases, showing the consequences of discounting regression analyses through the presentation of invalid or overweighed critiques. This analysis shows that when the defendant presents critiques of the plaintiff's regression, the plaintiff is statistically significantly less likely to prevail. This Article concludes by proposing that courts adopt a peer review system to evaluate the validity of critiques proffered by opposing counsel during evidentiary deliberations.

19. *See, e.g.*, E.E.O.C. v. Sears, Roebuck & Co., 628 F. Supp. 1264, 1344 (N.D. Ill. 1986) (discussing the court's finding that important variables were omitted) *aff'd* 839 F.2d 302 (7th Cir. 1988).

II. Econometrics in the Courtroom

As the Supreme Court acknowledged in *Teamsters*, regression analysis serves an important role in establishing discrimination.[20] As a result, parties often introduce regression analyses in Title VII employment discrimination cases.[21] One scholar noted in 1992 that "since [*Teamsters*], statistical evidence, most commonly multiple regression analysis, has become the primary means of establishing wage discrimination in disparate treatment cases."[22] Generally, the regressions help establish that the individuals were less likely to receive a promotion or to be hired or that they received lower wages because they were members of a protected class.[23] Of course, as was acknowledged in *Bazemore v. Friday*,[24] regression analyses, when flawed, can

20. *See* Int'l Bhd. of Teamsters v. United States, 431 U.S. 324, 339 (1977) (noting that statistical analyses play an important role in cases where discrimination is a disputed issue).

21. *See, e.g.*, Bazemore v. Friday, 478 U.S. 385, 399 (1986) (per curiam) (discussing the petitioners' heavy reliance on multiple regression analyses to demonstrate a pay disparity based on race); Morgan v. United Parcel Serv. of Am., 380 F.3d 459, 468 (8th Cir. 2004) (noting methodological deficiencies in a plaintiff's regression analyses in a Title VII case); Segar v. Smith, 738 F.2d 1249, 1261 (D.C. Cir. 1984) ("Multiple regression is a form of statistical analysis used increasingly in Title VII actions . . . ."). Plaintiffs also often present such evidence in Fair Housing Act claims, in RICO claims, and in constitutional challenges based on the Equal Protection Clause. *See, e.g.*, Rodrigues v. Nat'l City Bank, 277 F.R.D. 148, 155 (E.D. Pa. 2011), (discussing the use of regression analyses in a Fair Housing Act case) *aff'd*, 726 F.3d 372 (3d Cir. 2013); Neurontin Mktg. & Sales Practices Litig. v. Pfizer, Inc., 712 F.3d 21, 30–31 (1st Cir. 2013) (noting that the primary evidence in a RICO case was expert testimony describing a regression analysis); Eng'g Contractors Ass'n of S. Fla. v. Metro. Dade Cnty., 122 F.3d 895, 917 (11th Cir. 1997) (discussing the use of regression analyses in an Equal Protection Clause case).

22. James T. McKeown, *Statistics for Wage Discrimination Cases: Why the Statistical Models Used Cannot Prove or Disprove Sex Discrimination*, 67 IND. L.J. 633, 633 (1992).

23. *See infra* notes 34–39 and accompanying text (giving one example of such regression). Title VII prohibits discrimination on the basis of race, sex, color, national origin, and religion. 42 U.S.C. § 2000e-2 (2012). In addition, the Age Discrimination in Employment Act of 1967 prohibits discrimination on the basis of age. 29 U.S.C. §§ 621–34. The Americans with Disabilities Act of 1990 prohibits discrimination on the basis of disability. 42 U.S.C. §§ 12101–12213. Finally, the Equal Pay Act of 1963 prohibits wage disparities between men and women for equal work. 29 U.S.C. § 206(d).

24. 478 U.S. 385 (1986).

provide inadequate support for such cases.[25] This Part provides an overview of how regression analyses are used in employment discrimination cases and discusses the evidentiary standards that federal courts follow when addressing regression analyses as evidence.

A. Econometrics in Employment Discrimination Cases

Plaintiffs often present expert testimony and reports that include regression analyses to support a claim of employment discrimination.[26] Such employment discrimination claims include claims of sex, race, color, or national origin discrimination under Title VII, age discrimination claims under the Age Discrimination in Employment Act (ADEA),[27] sex discrimination under the Equal Pay Act (EPA),[28] and disability discrimination under the Americans with Disabilities Act (ADA).[29] In these cases, the plaintiffs' expert witnesses present statistics showing that, all other qualifications equal, being a member of a protected class decreased the plaintiff's expected wage or likelihood of receiving a promotion or being hired.[30] Alternatively, defendants often present regression analyses to establish that there was not a differential in hiring, promotions, or wages between the protected class and other similarly situated employees.[31]

While regression analyses are common in class action cases, such as *Wal-Mart Stores, Inc. v. Dukes*,[32] plaintiffs also often

---

25. *See id.* at 400 n.10 (1986) ("There may, of course, be some regressions so incomplete as to be inadmissible as irrelevant; but such was clearly not the case here.").

26. *See* cases cited *supra* note 21 (citing cases in which regression analyses were conducted by experts and presented as evidence).

27. 29 U.S.C. §§ 621–34.

28. *Id.* § 206(d).

29. 42 U.S.C. §§ 12101–12213 (2012).

30. *See, e.g.*, Lavin-McEleney v. Marist Coll., 239 F.3d 476, 478 (2d Cir. 2001) (discussing a statistical analysis that showed the plaintiff was paid less than male professors even after controlling for relevant factors such as experience, tenure status, and type of degree).

31. *See, e.g.*, Morgan v. United Parcel Serv. of Am., Inc., 143 F. Supp. 2d 1143, 1151 (E.D. Mo. 2000) (explaining the defendant's expert's use of statistical evidence to rebut the contention that a wage disparity was based on race).

32. *See* 131 S. Ct. 2541, 2555 (2011) (discussing regression analyses the

introduce regression analyses in individual employment discrimination claims.[33] In individual claims, this evidence can be used to establish disparate treatment claims, which allege that the employer treated the plaintiff worse than similarly situated individuals due to his or her protected class, or to establish underlying disparate impact claims, which allege that the defendant's policies have a differential impact on members of a protected class.[34] For example, in *Lavin-McEleney v. Marist College*,[35] the plaintiff, Ms. Lavin-McEleney, filed a disparate treatment claim, alleging that her employer, Marist College, paid her lower wages than her male counterparts.[36] To establish such a claim, the plaintiff presented expert-witness reports that included regression analyses, which analyzed the wages of each professor at Marist College.[37] These regressions controlled for characteristics that could influence each professor's wage separately from his or her sex, and the results showed a significant wage disparity on the basis of sex.[38] This evidence, along with anecdotal evidence, led the jury to find for the plaintiff and led the Second Circuit to uphold this decision.[39]

Plaintiffs often present regression analyses as evidence in class action discrimination cases to support a pattern or practice in a disparate treatment discrimination claim and to establish

---

plaintiffs argued were evidence of commonality).

33. *See* Derrickson v. Circuit City Stores, Inc., 84 F. Supp. 2d 679, 689 (D. Md. 2000) (noting the use of a regression analysis as evidence in an individual's claim that he was denied a promotion based on his race).

34. Kingsley R. Browne, *Statistical Proof of Discrimination: Beyond "Damned Lies*," 68 WASH. L. REV. 477, 481 (1993) ("Under disparate-impact theory, the plaintiff challenges a facially neutral employment practice on the ground that it produces an adverse—if only inadvertent—effect on a protected group.").

35. 239 F.3d 476 (2d Cir. 2001).

36. *See id.* at 478 (discussing the plaintiff's allegation that her raises were discriminatory because she was not promoted to a full professor despite her request to have her salary reviewed for gender disparity).

37. *See id.* at 482 (noting that the expert used salaries of the entire faculty to attain a sufficiently large sample size).

38. *See id.* at 478 ("[The plaintiff's expert] found that the plaintiff was paid significantly less than comparable male professors within the division.").

39. *See id.* at 481 ("We hold that statistical evidence of gender based salary disparity among comparable professors properly contributed to plaintiff's case in conjunction with her identification of a specific male comparator.").

commonality between the members of the class as required by statute.[40] Notably, the Supreme Court of the United States addressed the requirement of commonality in *Wal-Mart Stores, Inc. v. Dukes*, a nationwide class action of female employees alleging that Wal-Mart discriminated against females in their pay and promotion practices.[41] In *Dukes*, the plaintiffs were seeking both injunctive and declaratory relief.[42] To establish commonality and a prima facie case of gender discrimination, the plaintiffs presented expert reports using regression analyses to show that the plaintiffs, as females, received statistically significant lower wages and were less likely to receive promotions than their male counterparts.[43] Ultimately, the Court thought that the region-by-region regressions were insufficient to establish that the discrimination was typical of the employer's practices because it could not establish a uniform, store-by-store wage and promotion disparity.[44]

However, since *Dukes*, courts have permitted regression analyses as support for more limited class claims. In *Ellis v. Costco Wholesale Corp.*,[45] the Northern District of California distinguished a nationwide class of female employees alleging

---

40. Plaintiffs in a class action can also allege disparate treatment claims. Browne, *supra* note 34.

41. *See* 131 S. Ct. 2541, 2547 (2011) ("[T]he Court of Appeals approved the certification of a class comprising about one and a half million plaintiffs, current and former employees of petitioner Wal-Mart who allege that the discretion exercised by their local supervisors over pay and promotion matters violates Title VII by discriminating against women.").

42. *See id.* ("In addition to injunctive and declaratory relief, the plaintiffs seek an award of back pay.").

43. *See id.* at 2555 (explaining that, after the plaintiffs' expert conducted a regression analysis, he concluded that "there are statistically significant disparities between men and women at Wal-Mart . . . [and] these disparities . . . can only be explained by gender discrimination" (citation omitted)). The Court had to address Rule 23 of the Federal Rules of Civil Procedure, which requires that "the party opposing the class has acted or refused to act on grounds that apply generally to the class, so that final injunctive relief or corresponding declaratory relief is appropriate respecting the class as a whole." *Id.* (quoting FED. R. CIV. P. 23).

44. *See id.* ("A regional pay disparity, for example, may be attributable to only a small set of Wal-Mart stores, and cannot by itself establish the uniform, store-by-store disparity upon which the plaintiffs' theory of commonality depends.").

45. 285 F.R.D. 492 (N.D. Cal. 2012).

gender discrimination against their employer from the class in *Dukes*.[46] The court distinguished the class because of its smaller size, because it was limited to two positions with uniform job descriptions, and because it identified specific practices of the employer in one type of promotion.[47] As a result, the court did look to the regression analyses to establish commonality, and because the regression analyses established class-wide (and not localized) gender disparities, the court found commonality and certified the class.[48]

Ellis shows that even after *Dukes*, regression analyses can provide evidence of commonality in class action employment discrimination cases as well as establish a prima facie case of employment discrimination (either disparate impact or disparate treatment claims). However, as was the case in both *Ellis* and *Dukes*, such analyses are usually heavily scrutinized by the opposing party's conflicting expert testimony.[49] Unfortunately, despite the presence of evidentiary standards to help guide the court, judges and juries are not often equipped to analyze the strength of such conflicting testimony.

B. Economists as Experts

Generally, regression analyses must be ruled admissible under Rule 702 of the Federal Rules of Evidence, which allows an expert qualified by "knowledge, skill, experience, training, or

---

46. *See id.* at 509 ("As explained further below, the proposed classes in the instant case differ from that examined in *Dukes* in several material ways.").

47. *Id.*

48. *See id.* at 530 ("Plaintiffs' statistical evidence demonstrates classwide—as opposed to fragmented or localized—gender disparities supporting its contention that Defendant's classwide practices yield classwide effects.").

49. *See id.* at 521 (noting that defendant's expert conducted a statistical study and found no evidence of gender disparity); Wal-Mart Stores, Inc. v. Dukes, 131 S. Ct. 2541, 2553 (2011) ("The parties dispute whether Bielby's [the plaintiffs' expert] testimony even met the standards for the admission of expert testimony under Federal Rule of Civil Procedure 702 and our *Daubert* case . . . ."). For a discussion of common critiques of regression analyses in class action cases, including sample size issues, see William T. Bielby & Pamela Coukos, "*Statistical Dueling" with Unconventional Weapons: What Courts Should Know About Experts in Employment Discrimination Class Actions*, 56 EMORY L.J. 1563 (2007).

education" to testify and give opinions if: (1) the testimony will assist the trier of fact; (2) it is "based on sufficient facts or data;" (3) it is "the product of reliable principles and methods;" and (4) "the expert has reliably applied the principles and methods to the facts of the case."[50] In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[51] the U.S. Supreme Court interpreted Rule 702 to require the judge to exercise general gatekeeping functions and limit scientific and technical expert testimony based on whether it "can be (and has been) tested, whether it has been subjected to peer review and publication, its known or potential error rate and the existence of maintenance of standards controlling its operation, and whether it has attracted widespread acceptance within a relevant scientific community."[52]

Generally, a federal judge must determine whether to grant motions to strike expert testimony under Rule 702, and the judge must make this decision under *Daubert*. However, the vague language of Rule 702 and of the *Daubert* standard do not provide much guidance for this decision, and the judge must decide whether the theory or technique behind the scientific testimony meets *Daubert*'s requirements. Ultimately, this is a large burden, and "[a]ssessing these factors can be daunting for experts trained in science—judges and their clerks, as scientific laymen, will have even more trouble."[53]

The *Daubert* analysis is very important for the introduction of expert reports on regression analyses because of the complex nature of the studies and the ability of the studies to be manipulated. When experts present regression analyses as evidence of employment discrimination in Title VII cases, it is very important for the judge to take his or her gatekeeping function under *Daubert* very seriously. It is worthwhile to

---

50. FED. R. EVID. 702.

51. 509 U.S. 579 (1993).

52. *Id.* at 593–94.

53. Lawrence S. Pinsky, *The Use of Scientific Peer Review and Colloquia to Assist Judges in the Admissibility Gatekeeping Mandated by* Daubert, 34 HOUS. L. REV. 527, 543 (1997); *see also* Justin P. Murphy, *Expert Witnesses at Trial: Where Are the Ethics?*, 14 GEO. J. LEGAL ETHICS 217, 227 (2000) ("The determination of reliability can present a significant burden for trial court judges. Trial court judges are asked under rule 702 to be 'better equipped than an honestly-testifying expert to know whether the expert's opinion is reliable. That is an unlikely premise.'" (citation omitted)).

consider the incentives of parties to litigation to present empirical evidence and, especially, the incentive of the plaintiff. Parties are not obligated to present statistical evidence of discrimination. This is especially true in disparate treatment cases, where specific examples of discriminatory treatment are likely to be more persuasive than dry statistics.[54]

Given the upfront costs involved in hiring an economic expert to conduct regression analyses, as well as the ease (as we show *infra*) with which defendants can rebut valid statistical evidence by misleading or confusing jurors, plaintiffs should only be incentivized to present regression evidence when the statistical methodology utilized is consistent with professional standards. As a result, the general concerns with expert testimony may be diminished in the presentation of regressions presented by the plaintiffs, making the *Daubert* analysis less important. However, defendants still have incentives to present invalid attacks, and those attacks should also be scrutinized. This Article proposes that not only should the judge consider the reliability of the regressions presented in favor of the plaintiff, but the judge must also consider the reliability of the critiques that the defendant presents because the introduction of invalid attacks on regression analyses can negate the presentation of reliable evidence that suggests discrimination. Unfortunately, courts have adopted the defendant's attacks on the plaintiff's regression analyses in many cases, and the Supreme Court has acknowledged when this adoption is problematic.[55]

54. *See* Int'l Bhd. of Teamsters v. United States, 431 U.S. 324, 399 (1977) ("[T]his was not a case in which the Government relied on 'statistics alone.' The individuals who testified about their personal experiences with the company brought the cold numbers convincingly to life.").

55. *See* Bazemore v. Friday, 478 U.S. 385, 401 (1986) (per curiam) (finding that "the Court of Appeals failed utterly to examine the regression analyses in light of all the evidence in the record"). The Court reasoned that, "[w]hile the omission of variables from a regression analysis may render the analysis less probative than it otherwise might be, it can hardly be said, absent some other infirmity, that an analysis which accounts for the major factors 'must be considered unacceptable as evidence of discrimination.'" *Id.* at 400 (citation omitted). Accordingly, the Court ruled that "[n]ormally, failure to include variables will affect the analysis' probativeness, not its admissibility." *Id.*

C. The Court's Recognition of Potential Problems

Even before *Daubert* controlled the introduction of expert evidence under Rule 702, the U.S. Supreme Court noted the methodological concerns of statistics as evidence of employment discrimination. In *International Brotherhood of Teamsters v. United States*,[56] the United States presented statistical evidence to support their claim of race discrimination in pay and promotion practices.[57] After emphasizing the value of such evidence, the Court then cautioned "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 of the surrounding facts and circumstances."[58] Following *Teamsters*, legal scholars also began to acknowledge the potential manipulation and problems associated with econometrics in the courtroom, and expert witnesses began to present convincing, but often invalid, critiques of the opponent expert's analysis that surrounded the choice of variables controlled for in the regression.[59]

A highly visible example of valid statistical evidence being rebutted following *Teamsters* occurred in *E.E.O.C. v. Sears, Roebuck & Co.*[60] At the time, Sears was the second largest private employer of women in the United States.[61] In *Sears*, the Equal Employment Opportunity Commission (EEOC) brought a sex discrimination suit against Sears and supported that suit with regression analyses that showed a disparity between the hiring

56. 431 U.S. 324 (1977).

57. *See id.* at 399–400 (discussing case law that supports the use of statistical evidence to establish discrimination).

58. *Id.* at 340.

59. *See* Daniel L. Rubinfeld, *Econometrics in the Courtroom*, 85 COLUM. L. REV. 1048, 1095 (1985) (arguing that the expanded use of multiple regression techniques is accompanied by the possibility of their misuse). To avoid misuse, Rubinfeld recommended that expert testimony include whether results were sensitive to the choice of variables used in the regression model. *Id.*

60. 628 F. Supp. 1264 (N.D. Ill. 1986).

61. *See* Thomas Haskell & Sanford Levinson, *Academic Freedom and Expert Witnessing: Historians and the Sears Case*, 66 TEX. L. REV. 1629, 1641 (1988) (noting that during the period covered by the litigation Sears was the second largest employer of women outside of the federal government).

and paying of males and females.[62] However, the court discounted the regression analysis that showed that females were less likely to be hired into higher-paying commission sales jobs at Sears because of the "omission and inadequate coding of important variables."[63] These factors included "the applicant's interest in commission sales and in the product to be sold, . . . physical appearance, assertiveness, the ability to communicate, friendliness, and economic motivation."[64] Even though the court recognized that these factors were difficult to quantify, it noted that the absence of the factors meant that the plaintiff expert's analyses were entitled to less weight.[65] The court also accorded less weight to the regressions analyzing the salaries of the employees because the regressions did not control for several measurable variables including "veteran status, marital status and size of family, leaves of absence and college major" and unquantifiable variables, including "loyalty, dedication, and motivation."[66] The Northern District of Illinois incorrectly relied on the premise that "[i]t is important to include all variables that significantly influence the dependent variable."[67]

The notion that it is important to include all variables that may affect the dependent variable in a regression analysis attempting to prove employment discrimination had become so

62. *See Sears*, 628 F. Supp. at 1302–03 (discussing the court's criticism of the EEOC's statistical evidence).

63. *See id.* at 1302 (concluding the EEOC's analysis was flawed because of its "failure to include in its analysis many important factors that significantly affect the hiring process").

64. *Id.* at 1302–03. The plaintiff's expert, Dr. Siskin, introduced compensation regressions that accounted for the following variables: sex; time in present assignment; time in present assignment squared; additional time in checklist; additional time in checklist squared; additional time at company; additional time at company squared; territory of employee; job performance; whether employee was hired as a college trainee; whether the facility was located in an urban area; and education. *Id.* at 1339. It should be noted that a regression that takes into account these factors easily meets professional standards for publication in peer-reviewed economics journals. *Infra* Part III.A.

65. *See Sears*, 628 F. Supp. at 1303 n.34 ("The court recognizes that these factors are not easily quantified for purposes of a statistical analysis, and that data relating to these factors was generally not available to EEOC . . . . Therefore, Dr. Siskin's analyses are entitled to less weight to the extent they do not incorporate these factors.").

66. *Id.* at 1344–45.

67. *Id.* at 1287. This incorrect reliance will be explained in Part III.A, *infra*.

prominent in the 1980s that some courts began to recognize its misuse. *United States Department of Treasury v. Harris Trust and Savings Bank*[68] was an administrative proceeding in which the Department of Labor and the Department of Treasury alleged that Harris discriminated against women and minorities in violation of Executive Order 11246.[69] During the proceeding, the plaintiff's expert presented a regression analysis that controlled for education, school major, experience, and prior experience, and the defendant challenged the regression due to omitted variables.[70] The Administrative Law Judge (ALJ) then recognized that every regression excludes certain variables that may affect an employment decision and injected a very satirical but telling story in footnote thirty-six:

> The story is told about how detailed records were kept between 1900 and 1982 of the amount of krill estimated to have been eaten by all Antarctic mammals. A statistical whiz, with unlimited use of free computer time, compared these observations with both the gross national product of Lithuania in 1985 and the sale of liters of wine in Andorra in 1986. He found several direct correlations. He concluded that he could show that krill eaten was an absolute predictor for all sorts of phenomena if given appropriate access to a free computer. It is also told that he received large fees in many court cases by testifying about how krill eaten in Antarctic was the missing variable in the statistical analysis of one party or another in merger and discrimination matters. Luckily, no such presentation was made in this case and this "omitted" variable was not addressed.[71]

Contrary to the court in *Sears*, the ALJ then stated that, "while the weight given the evidence may be reduced as a refinement of the variables is made, [the U.S. expert's] study still contributes to the Plaintiff's case."[72]

68. 78-OFC-2, ALJ's Recommended Decision (Dep't of Labor Dec. 22, 1986).

69. *Id.* at 4.

70. *See id.* at 24 (noting that Harris attacked the validity of the government's statistical evidence by contending that adjustment bias and omitted variables permeated the statistical evidence).

71. *Id.* at 33 n.36.

72. *Id.* at 33.

The Supreme Court addressed in *Bazemore v. Friday*[73] the legitimacy of regression analyses in employment discrimination cases even when such regressions do not include every variable the defendant claims is relevant.[74] In *Bazemore*, multiple black employees alleged racial discrimination in payment practices.[75] To support this claim, the plaintiffs introduced statistical evidence, including regression results that showed a large pay disparity between black and white employees with the same job title, education, and tenure.[76] However, the District Court refused to accept the evidence as proof of discrimination, and the Court of Appeals upheld that determination.[77] The Supreme Court addressed the potentially valid reason for such refusal: the regressions failed to consider "a number of variable factors" that were relevant in salary considerations.[78] Although the regressions controlled for the variables that were identified by an Extension Service official as most determinative of salary (education, tenure, and job title) in addition to race, the defendants offered nine additional variables that they claimed needed to be included for the regression to be valid.[79] The defendant argued that the plaintiffs' failure to include these

73. 478 U.S. 385, 400 (1986).

74. *See id.* at 400 (per curiam) ("While the omission of variables from a regression analysis may render the analysis less probative than it otherwise might be, it can hardly be said, absent some other infirmity, that an analysis which accounts for the major factors 'must be considered unacceptable as evidence of discrimination.'").

75. *Id.* at 394.

76. *See id.* at 398 (discussing the variables used in the regression analysis and explaining that the "[p]etitioners selected these variables based on discovery testimony by an Extension Service official that four factors were determinative of salary: education, tenure, job title, and job performance"). The average pay disparity in 1975 was $395 a year, which was a disparity of about 3% of average annual salary in that year ($12,524). The average pay disparity in 1974 was $331 a year. *Id.* at 399.

77. *See id.* at 399 ("The Court of Appeals stated: [t]he district court refused to accept plaintiffs' expert testimony as proof of discrimination . . . because the plaintiffs' expert had not included a number of variable factors the court considered relevant . . . . The district court was, of course, correct in this analysis.").

78. *Id.*

79. *See id.* at 404 n.15 (noting that the district court listed nine variables it believed petitioners should have accounted for in their regression).

variables resulted in a false showing of discrimination.[80] But the Supreme Court recognized that, even though omitted variables can make regression analyses less probative, this consideration should not usually be made at the admissibility stage.[81] In fact, the Supreme Court noted that, because the burden of proof is preponderance of the evidence, regression analyses that do not include "all measurable variables" can "serve to prove a plaintiff's case."[82] As a result, the Court remanded the case for the lower court to consider the statistical evidence in light of the entire record.[83] Unfortunately, some courts and opposing experts still maintain that if any seemingly plausible variable can be declared an "omitted" variable, then the regression analysis is too unreliable to "prove a plaintiff's case."[84]

Following *Bazemore*, courts should have been less likely to discount the proof offered by regression analyses that fail to include every measureable variable. Unfortunately, *Bazemore* did not influence all courts in this manner. In fact, after *Bazemore*, the Seventh Circuit addressed the omitted variables in *Sears* and found "that the EEOC's failure to support its choice of variables in this case casts a shadow on the probative value of the regression analyses incorporating those variables."[85] The Seventh Circuit acknowledged *Bazemore* but recognized that the district court likely considered the regressions to be "so incomplete as to be inadmissible as irrelevant," which is the exception to the admissibility standards as recognized by the Supreme Court in *Bazemore*.[86] The Seventh Circuit found that the district court's

80. *See id.* at 399–400 (noting that the district court found that the regression analysis was not valid evidence of discrimination because experts failed to include variables which "ought to be reasonably viewed as determinants of salary").

81. *See id.* at 400 (finding that failure to include variables affects probability, not admissibility).

82. *Id.*

83. *See id.* at 386–87 (holding that the "Court of Appeals erred in disregarding petitioners' statistical analysis . . . [and] that on remand, the Court of Appeals should examine all of the evidence in the record . . .").

84. *See* E.E.O.C. v. Sears, Roebuck & Co., 839 F.2d 302, 349 (7th Cir. 1988) (finding that the district court did not err in concluding that the EEOC regression analysis was flawed due to omitted variables and incomplete data).

85. *Id.* at 326.

86. *Id.* at 327 (quoting Bazemore v. Friday, 478 U.S. 385, 400 n.10 (1986)). The Fourth Circuit also addressed how *Bazemore* applies during a summary

criticisms of the regression analysis were not clearly erroneous and upheld the decision for the defendant.[87]

The lower court decision in *Dukes v. Wal-Mart Stores, Inc.* provides an example of the court properly acknowledging *Bazemore*, but it also shows that defendants continued to make the same arguments following *Bazemore*. To establish class commonality and underlying disparate treatment in wages for women, the plaintiffs presented regression analyses that controlled for a number of major variables, including: "gender, length of time with the company, number of weeks worked during the year, whether the employee was hir[ed] or terminated during the year, full-time or part-time, which store the employee worked in, whether the employee was ever hired into a management position, job position, and job review ratings."[88] The defendant's expert (Dr. Haworth, who was also the expert in *Sears*) claimed, "[T]hese variables do not fully reflect [Wal-Mart's] compensation decision-making structure, thereby leaving open the possibility that one or more missing variables could explain the gender disparities in question."[89] The eleven other variables that Dr. Haworth recognized were quite similar to those she recognized in *Sears*: "hours worked, seniority, leave of absence, full-time/part-time status at hire, recent promotion or demotion, prior grocery experience, pay group, night shift, department, store size, and store profitability."[90] When the defendant sought to exclude the plaintiffs' regression from trial, the Northern District of

---

judgment motion and agreed that, due to the omission of variables measuring performance, the probative value of the regression analysis was an issue of material fact for the jury to decide in *Smith v. Virginia Commonwealth University*. *See* 84 F.3d 672, 676–77 (4th Cir. 1996) (finding that a question of material fact existed as to whether actual performance factors should have been included in the university's analysis). In this case, the defendant, VCU, presented a regression analysis to support its adoption of an affirmative action program for women, and the district court relied on this regression to find for VCU on summary judgment; however, due to the alleged flaws of the regression, the Fourth Circuit overturned the motion. *Id.* at 677.

87. *See Sears, Roebuck & Co.*, 839 F.2d at 348 (noting that "the district court's finding that the EEOC had not proved sex discrimination in wages through its flawed multiple regression analyses is not clearly erroneous").

88. Dukes v. Wal-Mart Stores, Inc., 222 F.R.D. 137, 159 (N.D. Cal 2004).

89. *Id.*

90. *Id.*; *see* E.E.O.C. v. Sears, Roebuck & Co., 839 F.2d 302, 330 (7th Cir. 1988) (discussing Haworth's regression analysis in *Sears*).

California quoted *Bazemore*, denied the motion, and determined that the regression went "well above the minimal threshold established by the courts, and thus his analysis is sufficient to raise an inference of discrimination for purposes of this motion."[91] Ultimately, the defendant continued to make this argument throughout the trial, and the argument influenced the Supreme Court when the Court rejected the regression analyses as proof of commonality or of disparate treatment.[92]

Many courts acknowledge *Bazemore* and the fact that econometric critiques challenging regression analyses should only affect the admissibility of regression analyses as evidence of employment discrimination in situations in which the regression analyses are "so incomplete as to be inadmissible as irrelevant."[93] However, defendants continue to make these arguments in court, and even though some courts correctly apply the admissibility standard, these arguments are also admitted, such that invalid critiques continue to diminish or eliminate the probativeness of the plaintiff's regression results. The three critiques that we found to be the most commonly argued in court are presented in the following section. The consequences of admitting the critiques when they are invalid are illustrated in the following empirical study. Recent examples of these consequences and of the court avoiding such consequences are presented in Part V.

III. Three Econometric Critiques

A review of employment discrimination judicial opinions and expert witness reports illustrates that opposing experts routinely offer the same three critiques to rebut a plaintiff's regression

---

91. *Dukes*, 222 F.R.D. at 160.

92. *See* Wal-Mart Stores, Inc. v. Dukes, 131 S. Ct. 2541, 2555–56 (2011) (finding that respondents' statistical proof and regression analyses failed to provide evidence of commonality either regionally, or, if the proof did exist, nationally).

93. *See Sears, Roebuck & Co.,* 839 F.2d at 326, 327 (7th Cir. 1988) (noting that the district court did not find the EEOC's analysis inadmissible due to failure to include variables, but instead found the analyses were not probative such that they were "so incomplete as to be inadmissible as irrelevant," qualifying as the exception to the rule in *Bazemore*) (quoting Bazemore v. Friday, 478 U.S. 385, 400 n.10 (1986)).

analysis. These three critiques are omitted variables, adequacy of sample, and lack of statistical significance.[94] To determine the prominence of these three issues and to analyze how courts treat each issue, we searched for all Title VII employment discrimination cases filed in the federal courts since 2000 that mention "regression analysis." This search resulted in a sample of seventy-eight cases.[95] This Part discusses these three prominent econometric issues as they pertain to Title VII employment discrimination cases and identifies when and why these critiques are overwhelmingly invalid.

A. Omitted Variables

In over 63% of the cases gathered, the court recognized that the opposing expert notes that the regression did not control for all measurable variables that may affect the treatment of the employees.[96] More specifically, courts frequently note (as the result of opposing expert testimony) that the plaintiff's regression analyses do not control for certain variables that the defendant argues are important determinants of employment decisions. This was the issue at hand in *Bazemore*.[97] Legal scholars have

94. *See generally* Bazemore v. Friday, 478 U.S. 385 (1986) (addressing omitted variables); Thomas v. Deloitte Consulting LP, No. 3-02-CV-0343-M, 2004 WL 1960097 (N.D. Tex. Sept. 2, 2004) (addressing issues relating to statistical significance); Coleman v. Exxon Chem. Corp., 162 F. Supp. 2d 593 (S.D. Tex. 2001) (noting problems with the sample size).

95. The citations to each case are listed in Table A of the Appendix. This sample is the result of a Westlaw search and, as such, does not represent all federal employment discrimination cases in which plaintiffs presented regression analyses. However, the statistics gathered from the search still provide insight into the prominence of such issues and anecdotal evidence gathered from the opinions provides insight into how courts address such issues. While the original search that resulted in the sample was for "regression analysis," a search for cases published between "regression analys!" results in the same sample. We did not expand the search to include "regression" because the courts almost always use "analysis[es]" and often use "regression" for its noneconometric meaning. The original search resulted in 177 cases; however, many of these cases were duplicates, many of the cases simply referenced other cases that presented regression analyses, and many of the cases simply referenced a Title VII case in the opinion.

96. *See infra* Table 2.

97. *See* 478 U.S. 385, 400–01 (1986) (noting the differences between plaintiffs' variables and defendants' variables).

recognized since 1986 that defendants attack plaintiff's regressions for "fail[ing] to account for important explanatory factors."[98] That all potential variables are not included is a characteristic of regression analysis and does not reflect shortcomings of the analysis. In fact, many personal characteristics, such as marital status and number of children that would be included in academic studies of earnings, are specifically excluded from earnings regressions in litigation because these personal characteristics are not legally relevant. Other reasons for excluding variables include that the variable at question may itself be a product of the discriminatory treatment at issue.[99]

The random error term, which is part of any regression equation, encompasses the effects of variables not directly included in the regression equation. Including more variables may result in higher explanatory power of the regression equation (what economists refer to as the *R* squared), but the

---

98. Barbara A. Norris, *Multiple Regression Analysis in Title VII Cases: A Structural Approach to Attacks of "Missing Factors" and "Pre-Act Discrimination,"* 49 L. & CONTEMP. PROBS. 63, 63 (1986).

99. *See* Trout v. Garrett, 780 F. Supp. 1396, 1412–13 (D.D.C. 1991)

> Regression analyses are typically challenged on the basis that one or more variables should be included or excluded because of their appropriateness or lack thereof. One basis for excluding a variable as "tainted" is that it gives a false explanation for the disparate impact. A prime example of tainted variables are "status variables," such as job rank or grade level, which could reflect, at least in part, prior discrimination. Baldus & Cole, Statistical Proof of Discrimination § 83 at 112–13 (1987 Supp.). If, for example, an individual's grade level is itself based on discrimination, then use of grade level as a variable would falsely suggest that disparities in pay were attributable to an objective factor rather than to the real source, discrimination. In this Circuit, it is the law that a variable is to be excluded if it is not demonstrated by clear, affirmative evidence that it is based on neutral, objective factors, applied consistently. Valentino v. United States Postal Service, 674 F.2d 56, 72 n.30 (D.C. Cir. 1982); *cf.* Sobel v. Yeshiva University, 839 F.2d 18 (2d Cir. 1988).

One scholar has argued that tainted variables can still assist a court in determining what type of decisions lead to disparities and in determining the appropriate level of damages. *See* Srijati Ananda & Kevin Gilmartin, *Inclusion of Potentially Tainted Variables in Regression Analyses for Employment Discrimination Cases*, 13 INDUS. REL. L.J. 121, 151 (1991) (noting that tainted variables can be helpful to achieve more accurate assessments of discrimination, which can have "obvious relevance in the shaping of appropriate injunctive and monetary relief").

only valid concern is whether failure to control for the alleged omitted variable causes "omitted variable bias." The practical consequence of omitted variable bias in employment discrimination cases is that the estimated regression equation erroneously shows discrimination when in fact the omitted factor is the legitimate and nondiscriminatory cause of the differential employment outcome between the protected class and the nonprotected class.[100] In addition to not being a "tainted variable," two conditions must hold for the omitted variables to cause the estimate of the coefficient of interest to be biased: the omitted variable must be correlated with the variable that represents the protected class at issue and the omitted variable must have a statistically significant effect on the outcome.[101] Often, the purported omitted variable will not have a statistically significant effect on the outcome.[102] The lack of an effect occurs because it is not important or because it is correlated with variables already included in the equation so that further inclusion of a related variable is redundant and adds little to the regression.[103]

For example, in *Bazemore*, if the primary omitted variable raised by the defendant (job performance) was correlated with the variable of interest (a variable indicating that the observation was a black individual), if performance rating had a statistically significant positive effect on pay, and if performance ratings were not themselves discriminatorily assigned, then omitted variable bias would result. If the omitted variable was negatively correlated with race (meaning black individuals have worse job performance), then its omission would bias the coefficient on the variable of interest upwards because job performance is positively correlated with wage; this bias would have meant that the significant positive coefficient on the variable of interest (black) may have been overstated. This bias would show a larger pay disparity due to race than would appear in a regression analysis

100. *See* JEFFREY M. WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH 90–91 (4th ed. 2009) (explaining omitted variable biases).

101. *See id.* at 91 (discussing the effect of the omitted variable bias on regression analyses).

102. *See id.* (noting that a small bias "need not be a cause for concern").

103. *See id.* at 91–92 (discussing the effect of sample size and variable correlation on a regression analysis).

that also controlled for job performance. Alternatively, if black individuals had better job performance, omission of job performance would show a lower disparity in pay on the basis of race.

In fact, in *Bazemore*, regressions that included job performance showed a larger race disparity than regressions that excluded this variable.[104] The remaining additional variables offered by the defendants to explain the observed pay disparity referred to county-to-county differences in salary increases.[105] However, unrebutted evidence showed that blacks were not disproportionately located in counties that contributed only a small amount to salary increases.[106] That is, the so-called omitted variables were not correlated with the variable of interest. Absent a correlation between the so-called omitted factors and the protected class, these omitted factors could not provide a race-neutral explanation for the pay disparity.[107] The above discussion is summarized in Takeaway One below.

> Takeaway One: An omitted variable that will only affect the results of a regression analysis establishes discrimination if the omitted variable is correlated with the variable of interest (likely an indicator variable for the individual being a member of the protected class) and is itself a statistically significant determinant of the outcome. Furthermore, many possible variables are legitimately excluded because they are not legally relevant, because they may themselves be the outcome of the discriminatory treatment at issue, or because they are adequately represented by variables already included in the regression equation.

---

104. *See* Bazemore v. Friday, 478 U.S. 385, 401 (1986). The pay disparity was $475 as compared to the values discussed. *Id.* The regressions were not presented at trial because performance ratings were missing from 20% of the employment records. *See id.* at 401 n.11 (noting missing data).

105. *Id.* at 404 n.15 (noting missing variables related to county-to-county differences).

106. *See id.* at 402 ("The United States presented evidence which it claims respondents did not rebut, establishing that black employees were not located disproportionately in the counties that contributed only a small amount to Extension Service salaries.").

107. *See id.* ("Absent a disproportionate concentration of blacks in such counties, it is difficult, if not impossible to understand how the fact that some counties contribute less to salaries than others could explain disparities between black and white salaries.").

B. Sample Size

In over 62% of the cases gathered, the court notes potential faults of the sample analyzed in the regression analyses. In these cases, the court either notes that the sample analyzed was not the sample that should have been analyzed, or the court notes that the sample is too small to draw certain conclusions from the case.[108] However, for a regression analysis to be statistically valid, the only requirement about sample size is that there are at least as many observations as parameters in the regression model.[109] Sample size affects the power of the estimates—the probability that a statistically significant effect, if true, can be detected with the given sample size.[110] Statistically significant results are less likely when the sample size is small.[111] As noted by Daniel Rubinfeld,

> Other things being equal, the statistical significance of a regression coefficient increases as the sample size increases. Thus, a $1 per hour wage differential between men and women that was determined to be insignificantly different from zero with a sample of 20 men and women could be highly significant if the sample were increased to 200.[112]

Valid conclusions can certainly be drawn from samples that are not very large, and finding statistically significant effects in

---

108. *Infra* Appendix A.

109. *See* WOOLDRIDGE, *supra* note 100, at 167

> For example, the unbiasedness of OLS (derived in Chapter 3) under the first four Gauss-Markov assumptions is a finite sample property because it holds for any sample size n (subject to the mild restriction that n must be at least as large as the total number of parameters in the regression model, k=1).

110. *See* Daniel L. Rubinfeld, *Reference Guide on Multiple Regression*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 179, 192 (Fed. Judicial Ctr., 2d ed. 2000) (noting that a difference could be "statistically significant" if a large enough sample is studied).

111. *See id.* (describing the possibility of obtaining results that are "practically significant, but statistically insignificant," particularly with small sample sizes).

112. *Id.* Notably, Rubinfeld uses a sample size of twenty as a comparison to a sample size of two hundred, which indicates that Rubinfeld considers a sample size of twenty to be acceptable for a regression analysis, subject only to the limitation that power is lower for a sample of twenty than for a sample of two hundred.

smaller sample sizes suggests that the estimated disparity is large, not that the estimates are invalid.[113] Confidence intervals and tests of statistical significance take into account the sample size and thus account for the greater variability in estimates that arise from smaller sample sizes relative to larger sample sizes.[114]

The unimportance of sample size is further demonstrated when the purpose of the regression analysis is not to draw conclusions beyond the sample or the employer under consideration. The regression analyses presented in employment discrimination cases are not meant to be representative of the entire U.S. population.[115] Instead, these regressions are only meant to establish whether that plaintiff was treated differently than similarly situated coworkers due to his or her protected class under a disparate treatment claim or that one of the defendant's policies had a disparate impact on members of a protected class.[116] Furthermore, the sample size is inherently limited by the number of employees in a firm or, in the case of discrimination in hiring, the records of applicants maintained by the firm.[117] Studies with similar goals and sample sizes as small as twenty observations have been the basis of articles published in reputable economic journals and often cited reports.[118] It is quite easy to find studies published in major economic journals

---

113. *See id.* at 191–92 (noting that even minor differences can be statistically significant if a sufficiently large sample size is studied).

114. *See id.* at 192 (explaining that statistical significance is partially determined by the sample size).

115. *See* Browne, *supra* note 34, at 506 (noting that comparisons between the employer's work force and the general population are not typically appropriate).

116. *See id.* at 478 (describing disparate treatment claims and disparate impact claims).

117. Note that Wal-Mart's "tap on the shoulder" approach made it impossible to assess whether promotions from within were representative of applicants. *See* Dukes v. Wal-Mart Stores, Inc., 222 F.R.D. 137, 148–49 (N.D. Cal. 2004) (describing the subjective factors involved in promotion and the "tap on the shoulder" approach).

118. *See* WOOLDRIDGE, *supra* note 100, at 167 (referencing regressions with samples sizes of twenty and seventy-two) (citing Ray C. Fair, *Econometrics and Presidential Elections*, 10 J. ECON. PERSP. 89, 98 (1996)); Leslie A. Whittington, James Alm & H. Elizabeth Peters, *Fertility and the Personal Exemption: Implicit Pronatalist Policy in the United States*, 80 AM. ECON. REV. 545, 545 (1990).

with sample sizes that would be deemed too "small" by an opposing expert in an employment discrimination case.[119] Unfortunately, opposing experts and judges often refute regression analyses due to sample sizes that are of similar or even far larger sizes based on nothing more than their assertion that larger samples are required and without articulating any scientific basis to support their claim.[120]

Reliable and strong conclusions can be drawn from small samples, especially when the studies do not draw externally valid conclusions, as is the case in the regression analyses presented in employment discrimination cases. As a result, as long as the plaintiff presents a regression analysis with a model that is properly specified, its admittance into the courtroom or the reliability of it should not be affected by sample size.

> Takeaway Two: Sample size affects only the statistical power and not the validity of the regression. Admissibility and reliability of regression evidence should not be based on sample size.

C. Statistical Significance

In close to 40% of the cases in our sample, the court notes a discrepancy in statistical significance. The court either notes that the defendant's and the plaintiff's experts drew conflicting conclusions about statistical significance or recognizes the lack of statistical significance of the variable of interest.[121] Not surprisingly, many courts require that results from a regression analysis be statistically significant to draw conclusions from

---

119. *See* Joni Hersch & W. Kip Viscusi, *Punitive Damages: How Judges and Juries Perform*, 33 J. LEGAL STUD. 1, 33 (2004) (reporting separate regressions for jury trials and bench trials based on 119 jury trials and 54 bench trials in Table Three); Joni Hersch, Alison F. Del Rossi & W. Kip Viscusi, *Voter Preferences and State Regulation of Smoking*, 42 ECON. INQUIRY 455, 464–66 (2004) (reporting regressions results in Tables Six, Seven, and Eight based on fifty states).

120. *See, e.g.*, Coleman v. Exxon Chem. Corp., 162 F. Supp. 2d. 593, 618–19 (S.D. Tex. 2001) (finding that, although the sample size was not too small as a matter of law, it was not sufficiently reliable to support the plaintiffs' claims of discrimination).

121. *Infra* Appendix A.

them.[122] Many courts also require statistical significance of the results for regression analyses to enter the courtroom under evidentiary standards.[123] But courts have demonstrated a fundamental confusion about what constitutes statistical significance, and this confusion is easily and frequently exploited, resulting in valid statistical evidence being deemed inadmissible.

There are three separate but related issues to consider in any determination of statistical significance. First, what level of significance is required? Second, should tests be one-sided or two-sided? Third, has the regression specification been manipulated to achieve a desired level of statistical significance?

Regarding the first issue, some courts adopt a bright-line rule regarding the admissibility and reliance of regression results due to statistical significance. These rules prevent the reliance on data not significant at the 5% level (which in a two-sided test is a p-value of .05 or approximately two standard deviations).[124] However, as noted by the Northern District of Texas in *Thomas v. Deloitte Consulting, L.P.*,[125] most courts, including the Fifth Circuit, have rejected such a bright-line standard.[126] Even though these courts reject a bright-line standard, the courts often prevent analyses with results that fall short of the two standard deviation requirement but are statistically significant at levels recognized in academic research from entering the courtroom under *Daubert*. In *Thomas*, the Northern District of Texas

---

122. *See* Thomas v. Deloitte Consulting L.P., No. 3-02-CV-0343-M, 2004 WL 1960097, at *4 (N.D. Tex. Sept. 2, 2004) (determining that a plaintiff may establish a prima facie case by using statistics if a gross disparity that is statistically significant is shown).

123. *See id.* at *5 (recognizing a bright-line rule for statistical significance that prevents evidence not significant at a 5% level from entering the courtroom).

124. *See id.* (citing several cases which established a bright-line rule of either 5% statistical significance or two standard deviations).

125. No. 3-02-CV-0343-M, 2004 WL 1960097 (N.D. Tex. Sept. 2, 2004).

126. *See id.* at *5 (recognizing that the Second, Third, Seventh, Fifth, and Eighth Circuits have rejected this bright-line standard and instead determine the statistical significance of a result on a case-by-case basis). The court also noted that the Fifth Circuit recognized in *Overton v. City of Austin*, 871 F.2d 529, 544 (5th Cir. 1989), that statistical significance is dependent on sample size, and the sample size varied with each analysis. *Id.* The Supreme Court has not addressed this issue, and "most courts agree that there is no bright-line test." 1 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (2011–2012).

recognized that "*Daubert* instructs that a court should consider the known or potential rate of error when assessing the scientific validity or reliability of expert testimony."[127] The court did not allow the plaintiff's expert to present regression results showing a gender pay disparity that ranged between 7% and 10% significance in a two-sided test of significance, which corresponds to statistical significance in a one-sided test of 5%.[128] The following passage illustrates the court's decision:

> The court is unaware of any employment case where the jury was allowed to consider statistical evidence of discrimination that approached the 10% level used by Dr. Sobol. To the contrary, "[s]tatisticians tend to discard chance as an explanation for a result when deviations from the expected value approach two standard deviations." *Payne v. Travenol Laboratories, Inc.,* 673 F.2d 798, 821 (5th Cir.), *cert. denied,* 459 U.S. 1038 (1982). Given the relatively small sample size used by Dr. Sobol, the court has little difficulty in concluding that a statistical deviation of 7% to 10% does not adequately rule out that the alleged disparities identified in her report were due to chance. As a result, Dr. Sobol will not be permitted to offer testimony regarding the results of her statistical analysis.[129]

The court supported this decision by citing several other Fifth Circuit cases that required 5% significance (making it seem as if the court applied a bright-line rule).[130] It is common practice in peer-reviewed research in economics to consider a result as "statistically significant" when the result is significant at the 10% level or less in a two-sided test.[131] If courts use *Daubert* to remove

---

127. *Thomas*, 2004 WL 1960097, at *5 (citing Daubert v. Merrell Dow Pharms., 509 U.S. 579, 594 (1993)).

128. *Id.*

129. *Id.*

130. *See id.* (citing E.E.O.C. v. Ethan Allan, Inc., 259 F. Supp. 2d 625, 635–36 (N.D. Ohio 2003) (requiring two standard deviations)); Elliot Grp. Med. & Surgical Serv., 714 F.2d 556, 565 (5th Cir. 1983), *cert. denied*, 467 U.S. 1215 (1984) (requiring a 5% level); Cooper v. Univ. of Tex. at Dall., 482 F. Supp. 187, 194 (N.D. Tex. 1979), *aff'd*, 648 F.2d 1039 (5th Cir. 1981) (requiring a 5% level).

131. Many studies published in major economic journals report results that are significant at the 10% level for a two-sided test and discuss these results as statistically significant. *See* Joni Hersch, *Home Production and Wages: Evidence from the American Time Use Survey*, 7 REV. ECON. HOUSEHOLD 159, 167 (2009) (indicating levels of significance at the 1%, 5% and 10% level); Joni Hersch & W. Kip Viscusi, *Immigrant Status and the Value of Statistical Life*, 45 J. HUM. RES.

results that are actually statistically significant, as the court did in *Thomas* and as the courts that adopt a bright-line 5% standard most certainly do, then valid and valuable evidence will not enter the courtroom.

Second, courts rarely note whether the level of statistical significance required is for a one-sided or two-sided hypothesis test. In two 1977 decisions, the Supreme Court introduced the notion that differences that correspond to "two or three standard deviations" are in some way meaningful in supporting an inference of discrimination.[132] Even this vague reference to "two or three standard deviations" reflects a fragile understanding of the meaning of statistical significance. There is a vast difference in the probability that a disparity of two standard deviations occurs by chance and the probability that a disparity of three standard deviations occurs by chance. Assuming we conduct two-sided tests in a large sample, the probability that a disparity of two standard deviations occurs by chance is 4.55%, but the probability that a disparity of three standard deviations occurs by chance is a mere 0.27%. In two-sided tests, the two standard deviation criterion corresponds roughly to the 5% significance level commonly accepted in statistics.[133] The three standard deviation criterion is well beyond a level of significance expected in statistics. In fact, even the more stringent 1% level of significance requires a standard deviation of only 2.56.[134]

In suggesting the two or three standard deviations criterion for statistical significance, the Supreme Court was silent on whether they anticipated the statistical tests to be one-sided, meaning the test for discrimination examined whether the protected class was treated worse than the nonprotected class, or

749, 758 (2010) (reporting results significant at the 5% and 10% level).

132. *See* Castaneda v. Partida, 430 U.S. 482, 496 n.17 (1977) ("As a general rule for such large samples, if the difference between the expected value and the observed number is greater than two or three standard deviations, then the hypothesis that the jury drawing was random would be suspect to a social scientist."); Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 311 n.17 (1977) ("Because a fluctuation of more than two or three standard deviations would undercut the hypothesis that decisions were being made randomly with respect to race . . . .").

133. The exact value in large samples is 1.96, not 2.

134. These statistics were calculated using Stata, based on the standard normal distribution.

two-sided, meaning the test for discrimination is simply that one party is treated differently than the other with no hypothesis about which party is preferred.[135] That cases of discrimination reach the courts with ambiguity about which party is the victim of alleged discrimination seems implausible, and some courts have recognized this absurdity.[136] The distinction between one-sided and two-sided tests is often crucial.[137] In a one-sided test, the 5% level of significance is reached with 1.645 standard deviations. In a two-sided test, the 5% level is reached with 1.96 standard deviations.[138] However, both the level of significance and whether the hypotheses tests must be one-sided or two-sided (also referred to as "one-tailed" and "two-tailed") determine whether any given result is "statistically significant."[139] As Daniel Rubinfeld writes in a federal court guide to regression analyses:

---

135. *See Castaneda*, 430 U.S. at 496 n.17 (identifying that a difference greater than two or three standard deviations is suspect, but not identifying whether the test was one-sided); *Hazelwood Sch. Dist.*, 433 U.S. at 311 n.17 (noting that fluctuations of more than two or three standard deviations do not support claims that decisions were made randomly, but failing to identify whether the tests were one-sided); *see also* Palmer v. Shultz, 815 F. 2d 84, 92 (D.C. Cir. 1987) (noting that the Supreme Court has not provided explicit guidance on the issue of one-tailed or two-tailed approaches).

136. *See Palmer*, 815 F.2d at 95 (D.C. Cir. 1987) (noting that claims of alleged discrimination involved complaints about both under- and over-selection, and that "statistically significant deviations in either direction from an equality in selection rates would constitute a prima facie case of unlawful discrimination" possibly leading to confusion about which party is the victim of discrimination).

137. *See* Rubinfeld, *Reference Guide*, *supra* note 110, at 195 (noting that the choice of either a one- or two-tailed test may affect an expert's acceptance or rejection of a null hypothesis).

138. This is explained clearly in *Palmer v. Schultz*:

> How can a 5% probability of randomness correspond both to a measurement of two standard deviations and a measurement of 1.65 standard deviations, one may reasonably ask? There is a legitimate answer: it depends on whether one is using a "one-tailed" or a "two-tailed" test of statistical significance. A disparity measuring 1.65 standard deviations corresponds to a 5% probability of randomness under a one-tailed test. A disparity measuring two standard deviations (to be more precise, 1.96 standard deviations) corresponds to a 5% probability of randomness under a two-tailed test.

Palmer, 815 F.2d at 92.

139. *See id.* at 93 (explaining that a number's statistical significance varies depending on whether a one-tailed or two-tailed test is used).

> When the expert evaluates the null hypothesis that a variable of interest has no association with a dependent variable against the alternative hypothesis that there is an association, a two-tailed test, which allows for the effect to be either positive or negative, is usually appropriate. A one-tailed test would usually be applied when the expert believes, perhaps on the basis of other direct evidence presented at trial, that the alternative hypothesis is either positive or negative, but not both. For example, an expert might use a one-tailed test in a patent infringement case if he or she strongly believes that the effect of the alleged infringement on the price of the infringed product was either zero or negative.[140]

The third issue that courts often misunderstand is that, by adding additional explanatory variables that may or may not be relevant, statistical significance can often easily be manipulated to tip the level of significance below any purported cutoff value. In *Cason v. Nissan Motor Acceptance Corp.*,[141] by adding in an additional sixty-seven variables to indicate month-year that a loan was made, the defendant was able to drive a p-value from .073 (statistically significant at the 5% level in a one-sided test) to .107, even though as a group, these additional sixty-seven month-year variables were statistically irrelevant.[142]

Courts also decrease the reliance of regression results when they misinterpret other measures of statistical significance. For example, in *Sears*, the Seventh Circuit correctly referred to z-values as the "number of standard deviations between the actual and expected figures."[143] The court, however, then referred to a z-value of 3.6 as "barely statistically significant" and a z-value of 2.9 as "less than statistically significant."[144] In reality, z-values of 3.6 and 2.9 are equivalent to p-values of less than .001 in a

---

140. *See* Rubinfeld, *Reference Guide*, *supra* note 110, at 194.

141. 28 F. App'x 392, 394 (6th Cir. 2002) (No. 00-6483) (on file with the Washington and Lee Law Review).

142. Expert Report for Plaintiff, Supplemental Report on Racial Impact of NMAC's Finance Charge Markup Policy at 45, Cason v. Nissan Motor Acceptance Corp., 28 F. App'x 392, 394 (6th Cir. 2002) (No. 00-6483) (on file with the Washington and Lee Law Review). We calculated the F-statistic from these reports to independently determine the additional variables were statistically insignificant.

143. E.E.O.C. v. Sears, Roebuck & Co., 839 F.2d 302, 323 (7th Cir. 1988).

144. *Id.* at 335–36.

standard normal distribution.[145] As a result, these z-values of 3.6 and 2.9 easily reach the standard for statistical significance (instead of being "barely statistically significant") at the 5% level in both one-sided and two-sided tests.[146]

One scholar has suggested that experts simply present p-values, instead of using the term "statistically significant," so that the jury can decide whether the statistical evidence is reliable.[147] This argument has some merit because the defendants would then be able to present evidence arguing that the level of significance is below any reasonable standard of reliability.[148] However, due to general concerns with the presentation of expert testimony (as discussed in Part II.B, *supra*), the court may want to use its gatekeeping role under *Daubert* to keep out results that do not meet the level of significance typically reported as meaningful in peer-reviewed publications—significance at the 10% level.

> Takeaway Three: Employment discrimination tests should always be one-sided tests and results that are significant at the 10% level should always be considered "statistically significant."

IV. A Statistical Analysis of Econometrics in the Courtroom

As illustrated above, regression analyses often provide critical evidence in employment discrimination claims, but the evidence can quickly be diminished by the opposing party's often

---

145. *See* DEP'T OF STATISTICS, TEX. A&M UNIV., STANDARD NORMAL DISTRIBUTION, https://www.stat.tamu.edu/~lzhou/stat302/standardnormaltable.pdf (listing p-values for a range of z-values).

146. *See Tests of Significance*, YALE UNIV. DEP'T OF STATISTICS, http://www.stat.yale.edu/Courses/1997-98/101/sigtest.htm (last visited Nov. 18, 2014) (describing significance levels for p-values at various percent levels) (on file with the Washington and Lee Law Review).

147. *See* D. H. Kaye, *Is Proof of Statistical Significance Relevant?*, 61 WASH. L. REV. 1333, 1339–40 (1986) ("As to the finding's admissibility, the issue is whether the testimony that the numbers are 'significant' sufficiently advances the understanding of the trier of fact to be worth the effort consumed in its presentation and explanation.").

148. *See id.* at 1344–45 (noting that statistical significance at the .05 level does not objectively prove a proposition as true and that "[s]tatistical evidence need not be dispositive to be helpful in building a prima facie case").

invalid critiques. While anecdotal and theoretical evidence of this problem is very persuasive, statistical evidence could inform the extent of the problem. We gathered a sample of employment discrimination cases in which one of the parties (generally the plaintiff) presented regression analyses with the hope of gaining more insight into the problem through sample statistics and our own regression analyses. We hoped to answer the following questions: (1) Do plaintiffs who present regression analyses in the sample of employment discrimination cases benefit from the evidence? (2) Is the value of the evidence diminished if the opposing party also presents regression analyses? (3) How often does the court acknowledge an opposing party's critiques of the regression analyses? And (4) does the acknowledgment of those critiques further negate the introduction of the analyses?

A. Data

To answer each of the above questions, we searched for all Title VII employment discrimination federal court decisions available on Westlaw since 2000 that mention "regression analysis." Specifically, we limited the Westlaw search to Title VII cases published between January 2000 and October 2013 containing the words "regression analysis." This search resulted in a sample of seventy-eight cases.[149] The citations to each case are listed in Table A of the Appendix. Because this sample was gathered from a Westlaw search, this analysis does not represent all employment discrimination cases in which plaintiffs presented regression analyses; however, we believe it still provides valuable information about how courts and juries address the introduction of regression analyses in employment discrimination cases. The total sample is comprised of summary judgment motions, evidentiary motions, trial verdicts, and both district court and court of appeals opinions.

After reading each decision, we coded the following characteristics of the case: whether it was a class action; the type of discrimination claim made; whether the EEOC represented the charging party; and whether a disparate impact claim was made.

149. *See supra* note 95 (discussing the Westlaw search and the resulting sample).

Most importantly, we noted the outcome of the motion or trial, and we noted which parties presented regression analyses supporting their claims. The result of the claim was coded as the result reported in the opinion being analyzed. We then coded whether the result was favorable for the plaintiff or the defendant. For example, if the motion to exclude the plaintiff's statistical evidence was denied, then the result was coded as being in favor of the plaintiff. On the other hand, if the motion was granted, then the result was coded as being in favor of the defendant.[150] Many of these evidentiary motions are not followed by trials with published opinions (as the case might have been settled or the opinion not published). As a result, the final outcome of the case is not necessarily the outcome that we analyzed.

B. General Summary Statistics

In a 1991 study, Catherine Connolly analyzed forty employment discrimination cases in which one of the parties presented regression analysis.[151] Connolly found that plaintiffs who presented regression analyses were most likely to prevail when both parties presented regression analyses, but that the plaintiff only prevailed 52.5% of the time.[152] Connolly also found that the plaintiffs did not receive a comparative advantage when they were the only party to submit regression analyses.[153] In addition, Connolly compared the plaintiff's highest chance of winning (52.5%) to previous estimations of a plaintiff's chance of prevailing in an employment discrimination case (or motion),

150. Occasionally, multiple motions are addressed in one opinion with some in favor of the defendant and some in favor of the plaintiff. These opinions were also coded as in favor of the plaintiff or the defendant. For example if both parties' motions to bar expert evidence were denied, and the class was certified, the opinion was coded as in the plaintiffs' favor.

151. *See* Catherine Connolly, *The Use of a Multiple Regression Analysis in Employment Discrimination Cases*, 10 POPULATION RES. & POL'Y REV. 117, 123 (1991) (noting that in twenty of those cases, both parties presented regression analysis; in twelve cases, only the plaintiff presented a regression; and in eight cases, only the defendant presented a regression).

152. *Id.*

153. *Id.*

which ranged from 31%–58%.[154] This comparison suggested that, as of 1991, presenting a regression analysis did not increase the plaintiff's likelihood of prevailing, which, assuming the regressions were valid, it should.[155] For a first look at our sample, we computed statistics similar to Connolly's statistics.

In our sample, the plaintiff presented a regression in all but four of the seventy-eight cases (94.87%). In addition, the plaintiff was the only party to present a regression in 51.28% of the cases, and the defendant was the only party to present a regression in only 5.13% of the cases.[156] These statistics are not surprising because, as Connolly recognized as well, the plaintiff has the burden to establish a prima facie case of discrimination and, in the absence of direct evidence, the plaintiff must submit circumstantial evidence (including statistical evidence) to establish a rebuttable prima facie case.[157] In fact, "[s]tatistical evidence is indispensable to a claim of disparate impact because the claim is that the challenged practice has an adverse effect on a group, not merely on an individual,"[158] and as a result, 66.67% of the sample presented a disparate impact claim. Also, as discussed in Part II, plaintiffs frequently (even after *Wal-Mart Stores, Inc. v. Dukes*) use regression analyses to establish class

---

154. *Id.* at 122 (citing Paul Burstein, *Attacking Sex Discrimination in the Labor Market: A Study in Law and Politics*, 67 SOC. FORCES 641, 657 (1989)). However, the plaintiff was at a disadvantage when the plaintiff did not present a regression and the defendant did. *Id.* at 123.

155. *See id.* at 122–23 (noting that plaintiffs had success between 31%–58% of the time with or without the use of regression analysis).

156. It was not clear in every reported opinion whether the defendant presented a regression analysis. If it was not clear that the defendant did not present a regression analysis, then it was assumed that the defendant did not present such results.

157. *See* Connolly, *supra* note 151, at 122 ("The more extensive use of regression analysis by plaintiffs may reflect the ordering of the burden of proof in an employment discrimination case."). Under *McDonnell Douglas Corp. v. Green*, the plaintiff generally must establish that: "(i) he belongs to a racial minority; (ii) he applied and was qualified for a job the employer was trying to fill; (iii) though qualified, he was rejected; and (iv) thereafter the employer continued to seek applicants with complainant's qualifications." 411 U.S. 792, 792–93 (1973). Plaintiffs often present statistical evidence to meet and strengthen requirement (iii); however, courts have recognized that statistical evidence alone is not enough, and, as a result, 94% of the sample also presented anecdotal evidence. Browne, *supra* note 34, at 481.

158. Browne, *supra* note 34, at 479.

commonality, and over 75% of our sample involved a class action.[159] Regression analyses are also the "core" of pattern or practice claims, which generally underlie class action cases.[160] Table 1 presents these summary statistics.

Table 1: Summary Statistics

Characteristic	**Percent of Cases with Characteristic**
Plaintiff Only Party to Present Regression	51.28% (40)
Defendant Only Party to Present Regression	5.13% (4)
Plaintiff & Defendant Present Regression	43.59% (34)
Class Action	76.92% (60)
EEOC Representation	6.41% (5)
Disparate Impact Claim Made	66.67% (52)
Disparate Impact Was the Only Claim Made	23.08% (18)

This sample was gathered from a Westlaw search limited to Title VII cases in which a party submitted a regression analysis. These cases were decided during January 2000–October 2013. The number of cases analyzed is seventy-eight. The number of cases with each characteristic is indicated in parentheses.

Nielsen et al. analyzed the outcome of a sample of employment discrimination cases filed in federal court during 1988 to 2003.[161] Their study reported that 6% of the employment discrimination cases they analyzed went to trial, and, of those 6%, the plaintiffs won 33% of the time.[162] In addition, of those cases that went to summary judgment, the plaintiffs prevailed approximately 43% of the time.[163] Nielsen et al.'s sample represents a broader sample than the sample analyzed in this Article, which is limited to cases in which regression analysis is presented. In our sample, the plaintiff wins 41.03% of the time.

159. *See id.* at 478–79 (describing the role of statistical evidence in a class action or pattern-or-practice case).

160. *See* Bell v. EPA, 232 F.3d 546, 553 (7th Cir. 2000) ("In a pattern and practice disparate treatment case, statistical evidence constitutes the core of a plaintiff's prima facie case.").

161. Laura Beth Nielsen, Robert L. Nelson & Ryon Lancaster, *Individual Justice or Collective Legal Mobilization? Employment Discrimination Litigation in the Post Civil Rights United States*, 7 J. EMPIRICAL LEGAL STUD. 175, 181 (2010).

162. *Id.* at 187.

163. *Id.* at 184.

Unlike in Connolly's sample,[164] in our sample, the plaintiff is most likely to win if they are the only party to present a regression (55%) and that percentage falls to 23.53% when the defendant also introduces a regression.[165] These percentages are presented in Table 2. When comparing these percentages to those presented in Nielsen et al., the plaintiff is actually at a comparative disadvantage when both parties present a regression analysis, as the plaintiff prevailed between 33%–43% in Nielsen's sample.[166] In addition, the plaintiff only gains a small comparative advantage when they are the only party to present a regression (55%).

Table 2: Summary Statistics of Plaintiff Result

Characteristic	**Percent of Cases with Plaintiff Result**
Plaintiff Only Party to Present Regression	55.00% (22/40)
Defendant Only Party to Present Regression	50.00% (2/4)
Plaintiff & Defendant Present Regression	23.53% (8/34)

This sample was gathered from a Westlaw search limited to Title VII cases in which a party submitted a regression analysis. These cases were decided during January 2000–October 2013. The number of cases analyzed is seventy-eight. The number of cases with plaintiff result out of the total number of cases with the characteristic is indicated in parentheses.

C. Statistical Findings Related to the Three Critiques

Assuming that the plaintiff presented a valid regression showing that the plaintiff was disadvantaged due to being a member of a protected class, the regression analysis should, on average, increase the plaintiff's probability of prevailing over the

164. Connolly, *supra* note 151, at 123.

165. Strangely, in our sample, the plaintiff wins 50% of the time when the defendant is the only party to present a regression. This sample is very small (n=4).

166. Nielsen et al., *supra* note 161, at 184. Our sample is not directly comparable to Nielsen et al. because this sample is limited to those with regression analysis and because the sample includes the outcomes of evidentiary motions in addition to the outcomes of trials and summary judgment motions. Our sample is also limited to reported cases.

defendant.[167] However, as evidenced by Connolly (1991) and by the updated analysis presented here, this increase does not actually occur.[168] This is likely because the defendant, even when not presenting a regression analysis, challenges the validity of the plaintiff's regression analysis by challenging the statistical methods used.[169] These challenges are usually based on the econometric critiques discussed above (omitted variables, sample size, and statistical significance).[170] In almost 90% of the cases analyzed, the court or opposing expert mentions at least one of these critiques.[171]

Table 3 presents summary statistics that show how often three econometric critiques are mentioned in the published opinions of our sample. These critiques were the three most common critiques mentioned by the court: omitted variables, inadequate sample, and a lack of statistical significance. As discussed above, these critiques are only valid in certain circumstances.[172] The statistics presented in Table 3 are limited to the cases where the plaintiff submitted a regression analysis, as our analysis focused on whether the plaintiff benefits by presenting such statistics in Title VII cases. We coded these

167. *See supra* Part II.B (discussing the incentives for the plaintiff to present a valid regression analysis).

168. *See* Connolly, *supra* note 151, at 122–23 (noting that the plaintiff has a similar probability of prevailing over the defendant with or without the use of regression analysis).

169. *See id.* at 123 ("Defendants . . . often successfully argue that the plaintiffs' computer print outs, mathematical equations, and university experts present a distorted view of the work environment. These defendants argue that personnel policies and practices are far too complicated to be reduced to a statistical showing.").

170. *See supra* Part III (analyzing three common econometric critiques).

171. *Infra* Table 3. If the court in any way referenced the regression analysis not including every relevant variable, we coded the opinion as referencing omitted variables. If the court in any way referenced the regression analysis not analyzing the correct sample or analyzing a sample that was too small, we coded the opinion as referencing critiques associated with the sample. If the court in any way mentioned that the regression results were not statistically significant or the fact that the opposing party challenged the level of significance, we coded the opinion as referencing statistical significance. Even if the court correctly analyzed these critiques, we still coded the court as referencing the critique in the opinion.

172. *See supra* Part III (discussing the validity of the three common econometric critiques).

critiques as present in an opinion regardless of whether the court or opposing expert correctly analyzed them.[173] Table 3 shows that the court discussed omitted variables in 63.51% of the cases. In addition, the court mentioned critiques associated with the sample (whether it was the correct sample or whether it was too small) in 62.16% of the cases, and the court mentioned critiques associated with statistical significance in 39.19% of the cases.

Table 3: Summary Statistics of Econometric Critiques

Econometric Critique	**Percent of Cases Presented**
Omitted Variables	63.51% (47)
Sample	62.16% (46)
Statistical Significance	39.19% (29)
Any Critique	89.19% (66)

This sample was gathered from a Westlaw search limited to Title VII cases. We exclude the four cases in which only the defendant presented a regression analysis. These cases were decided during January 2000–October 2013. The number of cases analyzed is seventy-four. The number of cases with the econometric critique is indicated in parentheses.

Table 4 presents summary statistics illustrating how often the plaintiff received a favorable result in opinions where the plaintiff presented a regression analysis and the court mentioned any of the three econometric critiques summarized in Table 3. As Table 4 illustrates, when the plaintiff presents regression results and any critique is mentioned, the plaintiff wins in 36.36% of the cases. This percentage is less than the percent of the total sample that wins when the plaintiff presents a regression analysis (40.54%).[174] In addition, the percentage of plaintiff verdicts is even smaller when omitted variables are discussed (31.91%).

173. In many cases, we could not discern whether the court correctly analyzed the critique or whether the critique was valid because we only had the published opinion available.

174. *See supra* Table 2 (obtaining this statistic from the total number of plaintiff results in this study, or 31/78).

Table 4: Summary Statistics of Plaintiff Result when Econometric Critique Discussed

Econometric Critique	Present of Cases With Plaintiff Result
Omitted Variables	31.91% (15/47)
Sample	39.13% (18/46)
Statistical Significance	37.93% (11/29)
Any Critique	36.36% (24/66)

This sample was gathered from a Westlaw search limited to Title VII cases. This data is limited to cases in which the plaintiff submitted a regression analysis. We exclude the four cases in which only the defendant presented a regression analysis. These cases were decided during January 2000–October 2013. The number of cases analyzed is seventy-four. The number of cases with plaintiff result out of the total number of cases with the econometric critique is indicated in parentheses.

Because contradictory statistics are not presented, a plaintiff should benefit most from presenting regression analyses when the plaintiff is the only party to present such statistical evidence. Table 2 illustrated that plaintiffs were more likely to receive a favorable result in cases with regression analyses when they were the only party to present such analyses. Table 5 presents summary statistics that show how often the court mentions econometric critiques even when the defendant does not present a regression analysis. As a result, Table 5 is limited to cases in which the plaintiff is the only party to present regression analyses. Comparing the second column of Table 5 to the statistics in Table 3 illustrates that the court is slightly less likely (82.50% compared to 89.19%) to discuss econometric critiques when the plaintiff is the only party to present a regression analysis. The third column of Table 5 shows that even when the plaintiff is the only party to present a regression analysis, the plaintiff has a smaller chance of winning when any critique is presented (48.48%) as compared to the total sample of plaintiffs that prevail when the plaintiff is the only party to present a regression analysis (55%). Not surprisingly, the percentages reported in column three of Table 5 are larger than those presented in Table 4 because the plaintiff is more likely to receive a favorable result when the plaintiff is the only party to present a regression analysis.

Table 5: Summary Statistics of Plaintiff Result when Plaintiff is the Only Party to Present Regressions and Critique is Presented

Critique	**Percent of Cases Presented**	**Present of Cases With Plaintiff Result**
Omitted Variables	57.50% (23)	43.48% (10/23)
Sample Size	55.00% (22)	50.00% (11/22)
Statistical Significance	40.00% (16)	37.50% (6/16)
Any Critique	82.50% (33)	48.48% (16/33)

This sample was gathered from a Westlaw search limited to Title VII cases. This data is limited to cases in which the plaintiff is the only party to submit a regression analysis. We exclude the four cases in which only the defendant presented a regression analysis. These cases were decided during January 2000–October 2013. The number of cases analyzed is forty. In the second column, the number of cases in which the critique is presented is reported in parentheses. In the third column, the number of cases with a plaintiff result out of the total number of cases in which the critique is presented is reported in parentheses.

D. Regression Results

To determine more accurately the consequences of an opposing party presenting contradicting regression analyses and critiques of the plaintiff's analysis, we conducted our own regression analyses. Each regression controls for major characteristics that we believe may affect the outcome of an employment discrimination case. In each analysis, the dependent variable is the outcome of the case or motion, and each regression controls for major characteristics that we believe may affect the outcome of an employment discrimination case. These characteristics are whether the case is a class action and whether the plaintiff was represented by the EEOC. We believe that plaintiffs in these cases may present more statistical evidence, which may affect the likelihood that the defendant challenges that evidence. In addition, the specifications control for whether the party presented only a disparate impact claim, as opposed to a disparate treatment claim or both claims, which may affect the likelihood that they prevail.

The variables of interest in our regression analysis are whether the defendant presented a regression and whether the reported opinion mentioned any of the discussed critiques. The dependent variable is whether the plaintiff received a favorable outcome, either at trial, from a summary judgment motion, from an evidentiary motion, or from a class action certification. Our regression analysis analyzes the seventy-four cases in which the plaintiff presented a regression analysis, as we are interested in the defendant challenging those regressions. The results of the ordinary least squares (OLS) regression are reported in Table 6 and show that, if the defendant presents a regression, the plaintiff is 28.8 percentage points less likely to have a favorable result. In addition, if the opinion mentions any of the econometric critiques (omitted variables, statistical significance, or sample deficiencies), then the plaintiff is 28.8 percentage points less likely to have a favorable result. Both of these results are significant at the 5% level in a two-sided test.[175]

175. As a result, this evidence should be submitted to a jury if presented in court.

Table 6: OLS Regression Results: Dependent Variable Plaintiff Result

Variable	**Coefficient**
Any Critique Present	-0.288** (0.117)
Defendant Presented Regression	-0.288** (0.136)
Class Action	0.191 (0.128)
EEOC	-0.465*** (0.134)
Disparate Impact Claim Only	-0.038 (0.140)
Constant	0.682*** (0.154)
Number of Observations	74

This sample was gathered from a Westlaw search limited to Title VII cases in which the plaintiff submitted a regression analysis. These cases were decided during January 2000–October 2013. Robust standard errors are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels in a two-sided test, respectively.

Table 7 presents results of a regression analysis that is limited to cases in which the plaintiff is the only party to submit regression analysis as evidence. Because of our sample construction, this regression does not control for whether the defendant presented a regression. These results show that the plaintiff is even more disadvantaged by critiques being discussed in an opinion when they are the only party to present a regression, as they are 36.0 percentage points less likely to receive a favorable result.

Although these results are limited because the sample is comprised only of cases and motions with published opinions, these results do show that the plaintiff is disadvantaged when econometric critiques, which may actually be flawed, are presented in court. Due to this strong result, this Article stresses the importance of the court exercising its gatekeeping role under *Daubert* in response to these critiques being presented. It also stresses the importance of the court and of the experts having an understanding of when these econometric critiques are actually invalid.

Table 7: OLS Regression Results for Cases Where Plaintiff is the Only Party to Present Regression Results: Dependent Variable Plaintiff Result

Variable	**Coefficient**
Any Critique Present	-0.360*** (0.111)
Class Action	0.208 (0.164)
EEOC	-0.703*** (0.227)
Disparate Impact Claim Only	-0.344 (0.206)
Constant	0.779*** (0.157)
Number of Observations	40

This sample was gathered from a Westlaw search limited to Title VII cases in which the plaintiff submitted a regression analysis. These cases were decided during January 2000–October 2013. Robust standard errors are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels in a two-sided test, respectively.

V. Examples of the Use and Misuse of Econometrics in Our Sample

The following examples from our sample of cases illustrate how the court has recently acknowledged the three critiques that were most often discussed in our sample. These examples show that the court is capable of correctly recognizing when the critiques are valid, but they also provide examples of how invalid critiques can lead the court astray.

A. Omitted Variables Examples

In *Sears*, the Northern District of Illinois completely misstated omitted variable bias: "However, the sex coefficient reflects not only the effect of sex, but also the residual effect of any factor which affects salary that is not included in the model. Thus, if important variables are omitted, the effect of sex on compensation estimated by the model will be artificially inflated."[176] Unfortunately, even after *Bazemore*, courts and opposing experts mischaracterize omitted variable bias and often do not focus on whether omitted variable *bias* is present in a regression. Instead, courts simply focus on the fact that variables that may explain part of the dependent variable are absent.[177] In addition, courts continue to allow defendants to present these arguments to the jury.[178] Within our sample of cases, there are examples of courts generally discussing omitted variables as a potential problem, as well as examples of the court incorrectly characterizing the problem. There are also examples of the courts correctly applying *Bazemore* and correctly recognizing when omitting certain variables is not an issue.

A more recent example of a court generally discussing the problem is the Eastern District of Pennsylvania's discussion in *Morgan v. United Parcel Service of America, Inc*.[179] To establish

176. E.E.O.C. v. Sears, Roebuck & Co., 628 F. Supp. 1264, 1344 (N.D. Ill. 1986).

177. *Id.*

178. *See, e.g.*, Dukes v. Wal-Mart Stores, Inc., 222 F.R.D. 137, 159 (N.D. Cal. 2004) (recognizing defendant's omitted variables critique).

179. 143 F. Supp. 2d 1143, 1151 (E.D. Mo. 2000).

race discrimination in wages, the plaintiff's expert presented a regression that controlled for the previous two performance evaluations but did not control for every performance evaluation.[180] When the opposing expert controlled for each evaluation in the regression, the coefficient on the variable indicating that the individual was black became insignificant.[181] The large increase in the number of variables within the equation (every evaluation instead of two) alone can result in lower statistical significance of explanatory variables.[182] However, this alternative explanation was never presented to the judge and never mentioned in the opinion. Even though the opposing expert did not show that these additional evaluations were negatively correlated with race and positively correlated with wage, nor did the expert show that taken as a group the additional explanatory variables resulted in a statistically significant improvement in explanatory power rather than merely a successful ruse to eliminate statistical significance in the original regression, the court concluded that the additional variables should be included in the regression and that the wage disparity between black and white managers was due to factors other than race.[183]

There are additional examples in our sample where a court accepted the defendant's argument that the plaintiff's regression is flawed due to omitted variables without requiring that the defendant establish the relationships discussed in Takeaway One.[184] In *Carpenter v. Boeing Co.*,[185] the defendant, Boeing, argued that the plaintiff's study did not "show that the 'something' causing men to work more overtime than women is

---

180. *Id.*

181. *Id.*

182. *See supra* notes 141–42 and accompanying text (noting that increasing the number of variables can manipulate the statistical significance).

183. *See Morgan*, 143 F. Supp. 2d at 1151–52 (finding that the defendant's expert properly included the additional variables and noting that the plaintiffs presented insufficient evidence to provide a basis for excluding the additional variables).

184. *See, e.g.*, Franklin v. Local 2 of the Sheet Metal Workers Int'l Ass'n, 565 F.3d 508, 518 (8th Cir. 2009) (finding the plaintiffs' statistical analysis unreliable because of omitted variables); Carpenter v. Boeing Co., 456 F.3d 1183, 1203–04 (10th Cir. 2006) (discussing the omitted variable as a flaw in the plaintiffs' regression analysis).

185. 456 F.3d 1183 (10th Cir. 2006).

the manager discretion that Plaintiffs have identified as the challenged employment practice" due to an omitted variable.[186] While the Tenth Circuit discussed the relationships that the omitted variable (department assignment) had with the outcome (overtime hours) and with the variable of interest (the protected class, in this case, female), it did not require the defendant to prove those relationships statistically or to prove that the inclusion of the variable affected the significance of the result.[187] The court upheld the lower court's decision denying class certification, in part, because of the flaws associated with the statistical analysis.[188]

In *Franklin v. Local 2 of the Sheet Metal Workers International Association*,[189] the Eighth Circuit reviewed the lower court's holding that "Dr. Gutman's report [was] not reliable because of the assumptions he m[ade], unsupported conclusions he dr[ew], and variables he fail[ed] to consider in rendering his opinion."[190] While the Eighth Circuit did discuss certain important relationships with the claimed omitted variables, it also quoted an earlier decision:

> The burden is on the opposing party to clearly rebut statistical evidence; hypotheses or conjecture will not suffice. When a plaintiff submits accurate statistical data, and a defendant alleges that relevant variables are excluded, defendant may not rely on hypothesis to lessen the probative value of plaintiff's statistical proof. Rather, defendant, in his rebuttal presentation, must either rework plaintiff's statistics incorporating the omitted factors or present other proof undermining plaintiff's claims.[191]

While this quote may seem to be in line with Takeaway One, as it requires some form of statistical proof that the variables are relevant, reworking the regression with the omitted variables

186. *Id.* at 1196.

187. *See id.* at 1195–96 (analyzing the experts' statistical findings).

188. *See id.* at 1203–04 (discussing the court's reasoning for finding the plaintiffs' statistical analysis deficient).

189. 565 F.3d 508 (8th Cir. 2009).

190. *Id.* at 514 (alternations in original) (quoting Franklin v. Sheet Metal Workers Int'l Ass'n Local Union No. 2, No. 06-0004-CV-W-GAF, 2008 WL 2819372, at *3 (W.D. Mo. July 8, 2008)).

191. *Id.* at 517 (quoting Coble v. Hot Springs Sch. Dist. No. 6, 682 F.2d 721, 730 (8th Cir. 1982)).

does not establish each of the required relationships; even if the variable of interest is no longer significant, the relationship with the omitted variable and the variable of interest is not proven.[192] In addition, an increase in the number of variables within the equation alone can result in lower statistical significance of explanatory variables.[193]

Fortunately, there are also examples in our sample of cases correctly applying *Bazemore* and not allowing claims of omitted variables to preclude the introduction of valid statistical evidence in employment discrimination cases. In *Derrickson v. Circuit City Stores, Inc.*,[194] the District of Maryland denied the defendant's motions for summary judgment and to exclude the plaintiff's expert report.[195] The plaintiff's expert report included results of a regression analysis that showed statistical disparities in promotion rates because of race.[196] The defendants sought to exclude the report, arguing that the regressions were flawed because they failed to control for store location.[197] The court then correctly cited *Bazemore* and denied the motion to exclude.[198] In fact, the court also correctly recognized that the plaintiff's expert did include store location in some regressions and found promotional differences that were statistically insignificant only due to sample size.[199] This discussion showed that the court understood the elements of Takeaway One and Takeaway Two, discussed below.

---

192. *See supra* note 101 and accompanying text (discussing the required relationships).

193. *See supra* notes 141–42 and accompanying text (noting that increasing the number of variables can manipulate the statistical significance).

194. 84 F. Supp. 2d 679 (D. Md. 2000).

195. *See id.* at 689–90 (denying the motions related to the defendant's challenge of statistical evidence offered to demonstrate employment discrimination).

196. *Id.* at 689.

197. *Id.*

198. *See id.* at 689–90 (noting that omission of a variable does not automatically render "an analysis which accounts for the major factors . . . unacceptable as evidence of discrimination" (quoting Bazemore v. Friday, 478 U.S. 385, 402 (1986))).

199. *See id.* at 690 ("Finally, Dr. Medoff did run the regression analysis to include store location as a variable and still found promotional differences to exist favoring whites. However, because the location variable reduced the sample sizes, many of the results were statistically insignificant.").

In *Lavin-McEleney v. Marist College*,[200] the Second Circuit also got it right when a defendant presented similar challenges to the plaintiff's statistical evidence.[201] In *Marist College*, the plaintiff presented regression analyses to support a claim of sex discrimination in wages.[202] These regressions controlled for characteristics that could influence each professor's wage separately from his or her sex.[203] These characteristics included each professor's rank, years of service, division, tenure status, and degrees earned.[204] Even after controlling for these variables, the coefficient for female was negative and significant, indicating statistically significant lower salaries for female employees.[205] As the Second Circuit recognized, the lower court properly admitted this statistical evidence despite the defendant's objections.[206] Also, despite the defendant's expert's contention that counterparts should only be compared on a departmental basis, the plaintiff's results were presented to the jury (as were the defendant's results that showed an insignificant gender-pay disparity).[207] Ultimately, the jury found that this evidence and additional anecdotal evidence supported a valid claim under the Equal Pay Act.[208] As a result, the district court awarded the plaintiff back pay and attorney's fees.[209]

---

200. 239 F.3d 476 (2d Cir. 2001).

201. *See id.* at 478–79 (explaining the defendant's objection to the plaintiff's statistical findings).

202. *Id.* at 478.

203. *Id.*

204. *Id.*

205. *Id.*

206. *See id.* at 482 (holding that the plaintiff's regression analysis "properly supported plaintiff's case and was appropriately employed to calculate damages").

207. *Id.* at 478–79.

208. *See id.* at 479 ("The jury found for the plaintiff on the Equal Pay Act claim, but decided that Marist's violation of the Act was not willful."). The special verdict form "instructed the jury not to consider plaintiff's Title VII violation if it found that Marist's violation of the Equal Pay Act was not willful." *Id.* Accordingly, "the jury did not find Marist liable on plaintiff's Title VII claim." *Id.*

209. *See id.* (noting the district court's decision to amend the judgment in the plaintiff's favor and award her back pay, attorney's fees, liquidated damages, and costs).

In *Tabor v. Hilti, Inc.*,[210] the most recent case in our sample, the court correctly analyzed omitted variable bias and almost directly addressed the points discussed in Takeaway One.[211] Although the Northern District of Oklahoma ultimately found for the defendant on the disparate impact claim of gender discrimination, the court correctly rejected the defendant's arguments challenging the plaintiff's regression analyses.[212] The court concluded as a matter of law that the failure to include priority ratings in the regression analysis did not render the regressions unreliable or unsound because the regressions "controlled for important variables other than sex that could impact promotion rates."[213] In fact, the court followed that finding with an even more detailed conclusion:

> Dr. Killingsworth's decision to not control for SMD mobility ratings does not render his analysis unreliable. In a regression analysis, mobility preferences would only change the statistical significance of the sex variable if mobility preferences differed by sex. However, Hilti provides no trustworthy data demonstrating that the mobility preferences of women differ from men among Base Market employees. Because the court may not presume such differences, the failure to control for mobility preferences does not make Dr. Killingsworth's analysis unreliable.[214]

This conclusion almost directly restates Takeaway One, showing that perhaps some courts are aware of the false critiques that expert witnesses present when attempting to impugn the reliability of valid statistical evidence presented by plaintiffs.

While it remains routine for defendants to attempt to refute regression analyses by claiming that omitted variables cause the illegal disparities, experts and judges must remember (as some courts in our sample have) that omitting variables that are expected to affect the dependent variable does not always lead to

---

210. No. 09-cv-189-GKF-PJC, 2013 WL 4068781 (N.D. Okla. Aug. 12, 2013).

211. *See id.* at *9 (analyzing the plaintiff's expert's findings and noting that the omitted variable "does not render his analysis unreliable").

212. *See id.* at *9, *11 (characterizing the plaintiff's expert report as "methodologically sound and reliable," but ultimately finding that the plaintiff did not meet her burden of proof).

213. *Id.* at *9.

214. *Id.* (citation omitted).

omitted variable bias and should not always negate the plaintiff's expert's regression results.

B. Sample Size Examples

Our sample of recent Title VII cases includes several cases in which the court discounted statistical evidence due to the size of the sample analyzed. In *Coleman v. Exxon Chemical Corp.*,[215] the Southern District of Texas noted that "[w]hether a sample is too small to yield meaningful results is a determination made by the district court on a case-by-case basis."[216] It also recognized that the Fifth Circuit had cautioned against relying on studies with small sample sizes.[217] In this race and gender discrimination case, the court held that the sample size of forty individuals (of which eight belonged to the protected class) was not inconclusive as a matter of law, but "any statistical analysis derived from such a small universe is far from conclusive and must be subjected to close scrutiny for reliability."[218] As a result, the court held that the statistical analysis was inadmissible.[219] As noted above, valid conclusions can be drawn from such a sample size. Unfortunately, although the court did not hold the evidence inconclusive as a matter of law, it still discounted the regression analysis and granted the defendant's motion for summary judgment.[220]

In *Guerrero v. Reno*,[221] the Northern District of Illinois addressed the defendant's motion for summary judgment on a disparate impact claim of national origin discrimination.[222] The

215. 162 F. Supp. 2d 593 (S.D. Tex. 2001).

216. *Id.* at 618 (quoting Anderson v. Douglas & Lomason Co., 26 F.3d 1277, 1289 n.20 (5th Cir. 1994)).

217. *See id.* (discussing the problems associated with small sample sizes).

218. *Id.* Part of this consideration was motivated by the fact that the inclusion of one outlier affected the results of the study. *Id.* at 618 n.34.

219. *See id.* at 617–20 (analyzing the plaintiffs' statistical analysis and finding "serious methodological flaws").

220. *See id.* at 618, 620–21 (summarizing the court's conclusions regarding the plaintiffs' statistical evidence).

221. No. 98 C 864, 2000 WL 1100400 (N.D. Ill. Aug. 7, 2000), *aff'd sub nom.* Guerrero v. Ashcroft, 253 F.3d 309 (7th Cir. 2001).

222. *Id.* at *1.

plaintiff presented a regression analysis to support his claim.[223] The regression showed a statistically significant disparity in the hiring of Hispanics for a specific job.[224] However, the defendant's expert attacked the report by arguing that a "sample size, of only thirty-four openings, was too small for reliable analysis."[225] Although the court did not explicitly state that this argument had merit, it did not give any weight to the regression analysis when determining that the plaintiff had not introduced enough evidence to survive summary judgment.[226]

In *Thomas v. Deloitte Consulting LP*, the defendant filed a motion to exclude the report of an expert statistician, which included a regression analysis.[227] The report was submitted to advance the plaintiff's claims of age and gender discrimination in the plaintiff's termination.[228] Ultimately, the court excluded the report due to concerns about statistical significance,[229] and this discussion is expanded on in the following section.[230] The court's final decision—that the statistical significance of the results made the results unreliable—was also based on the "relatively small sample size."[231] However, as noted above, smaller samples actually make it more difficult to find statistically significant results, and thus, the court's statement was misguided. In fact, the small sample size should have led the court to be more accepting of higher levels of statistical significance, and as a result, the court likely should not have excluded this evidence on this basis.

---

223. *Id.* at *6.

224. *Id.*

225. *Id.* at *7.

226. *See id.* at *14 (discussing the court's reasoning for finding insufficient evidence to support the plaintiff's claim).

227. No. 3–02–CV–0343–M, 2004 WL 1960097, at *1, *3 (N.D. Tex. Sept. 2, 2004).

228. *Id* at *1.

229. *See id.* at *6 (excluding the plaintiff's report because the analysis "either failed to test for statistical significance or did not use the proper threshold for statistical significance").

230. *See infra* Part V.C (citing examples of cases in which courts addressed statistical significance issues).

231. *See Thomas*, 2004 WL 1960097, at *5 (discussing the court's reasoning for excluding the plaintiff's expert report).

The above examples from the sample of cases that we analyzed illustrate how arguments of sample size can taint a court's decision at stages as early as evidentiary and summary judgment motions. Even if these arguments do not lead to an exclusionary motion or summary judgment ruling for the defendant, if the arguments are made again in the courtroom, they still have the opportunity to influence the judge or jury.

C. Statistical Significance Examples

In addition to *Thomas*, discussed in Part III.B above, in our sample of cases, there are several examples of the court strictly requiring a certain level of statistical significance for regression results to be admissible and persuasive. In *E.E.O.C. v. Autozone, Inc.*,[232] the EEOC brought a disparate treatment claim of race discrimination and a pattern or practice claim of gender discrimination.[233] To support the sex discrimination claim, the EEOC presented a regression analysis; however, the defendant challenged the analysis, claiming that the statistical significance of the main result was not reliable because it was significant at the 5% level and the Supreme Court had previously required significance at 2.3% in a different case.[234] Even though the court correctly identified that the previous case dealt with a one-sided test, and this case dealt with a two-sided test, the court still implied that it would require 5% significance in a two-sided test.[235] The court stated that "an approximation of two standard deviations at 5% is acceptable."[236] Because the plaintiff's results were significant at 5% in a two-sided test, the court did rely on

---

232. No. 00-2923 Ma/A, 2006 WL 2524093 (W.D. Tenn. Aug. 29, 2006).

233. *Id.* at *1.

234. *See id.* at *3 (criticizing the plaintiff's regression analysis because it used "an arbitrary significance level that [did] not conform to the requirements of *Castaneda v. Partida*" (citing Castaneda v. Partida, 430 U.S. 482, 496 n.17 (1977))).

235. *See id.* (noting that "[t]wo standard deviations is often approximated at 5% for two-tailed tests" (citing Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 318 n.5 (1977) (Stevens, J., dissenting))).

236. *Id.* (citing Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 311 n.17 (1977)).

the statistical evidence and did not grant summary judgment on the sex discrimination claim on that basis.[237]

In *Boyd v. Interstate Brands Corps.*,[238] the plaintiffs presented regression analyses to support class certification for their race discrimination claims.[239] These regressions sought to prove a statistically significant disparity in promotions based on race.[240] However, the court and the opposing experts challenged the results because they were not statistically significant.[241] The plaintiff's expert's report found results that were statistically significant at the 7% level (or with p-values of .07) in a two-sided test.[242] As a result, the report found that "the disparity in promotions for the relevant period was 'within 0.02 of being statistically significant.'"[243] Unfortunately, because of this (incorrect) statement, the court held that it did not even need to address the credibility of the report to determine that the plaintiffs did not meet their burden in establishing commonality.[244] If the Eastern District of New York had not previously required statistical significance at the 5% level, then the expert would not have likely stated that his results were not statistically significant; perhaps, the court should have ignored the expert's statement and relied on this valid statistical evidence to show commonality. If courts continue to apply such strict bright-line standards, then valid statistical evidence will not be introduced to support the claims of employment discrimination.

237. *See id.* at *3–7 (rejecting the defendant's argument that the plaintiff used an arbitrary significance level, but ultimately deciding that the plaintiff's results could not be considered relevant evidence due to flaws in the regression analyses).

238. 256 F.R.D. 340 (E.D.N.Y. 2009).

239. *Id.* at 362.

240. *Id.* at 360.

241. *See id.* at 361 (noting that the plaintiff's expert "did not find a statistically significant disparity in promotion rates between African-American and non-African-American employees").

242. *See* Declaration of Mark R. Killingsworth ¶ 15, tbl. 5, Boyd v. Interstate Brands Corps., 256 F.R.D. 340 (E.D.N.Y. 2009) (No. 00-CV-2249 (JFB)(RML)), ECF No. 132-18 (summarizing the expert's statistical findings regarding promotions and race).

243. *Boyd*, 256 F.R.D. at 362.

244. *See id.* (determining that "plaintiffs cannot get past the fact that Dr. Killingsworth did not find a statistically significant disparity in promotion rates").

VI. Potential Solution

Regression analyses can provide valuable evidence for both parties in employment discrimination cases, where direct evidence is often hard to come by. While we recognize the ability of experts to manipulate statistical evidence and the unreliability of certain techniques, we also recognize that these downfalls only occur in very limited circumstances. Opposing counsel and their experts are expected to attack the introduction of any evidence, but these attacks can also be manipulated and unreliable. When a plaintiff presents regression results establishing that she was treated differently in the workplace because she was a member of the protected class, the defendant often presents regression results contradicting those results. In addition, the defendant critiques the plaintiff's regressions. However, three of the most common arguments made (that the regression suffers from omitted variables, a small sample size, and a lack of statistical significance) are only arguments with true merit in very few circumstances.[245] As illustrated in Part IV.D, the introduction of these arguments decreases the probability that the plaintiff prevails and decreases the significance of presenting valid regression results that support the plaintiff's case. As a result, this Article proposes that the court exercise its gatekeeping function by either acting under *Daubert* or establishing a peer-review system to guarantee that only valid challenges to regression results enter the courtroom.

A. Using Daubert

Although it is a difficult task, judges are instructed under *Daubert* to consider whether expert testimony "can be (and has been) tested, whether it has been subjected to peer review and publication, its known or potential error rate and the existence of maintenance of standards controlling its operation, and whether it has attracted widespread acceptance within a relevant scientific community" before allowing the testimony to enter the courtroom under Federal Rules of Evidence 702.[246] As a result,

245. *Supra* Part III.

246. Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 580 (1993); *see*

judges will likely analyze whether expert reports that present regression results to establish or to refute employment discrimination meet the *Daubert* considerations.[247] However, it is also important for judges to consider whether the plaintiff's expert's attacks of the opposing expert's statistical techniques also meet the standards of *Daubert*.

It is just as important that judges attempt to determine whether these criticisms are valid because they too have the ability to persuade the jury; unjust criticisms can persuade the jury to reject valid statistical evidence that can assist the plaintiff in a discrimination case. Unfortunately, judges may not be aware of the takeaways presented above and may be unable to determine whether certain econometric critiques are actually invalid. This Article proposes that judges consider these takeaways and remember to analyze the reliability of criticisms found in expert reports instead of only analyzing the actual regression analyses.

Of course, we acknowledge that for judges to accurately make this decision they must be at least familiar with these three criticisms. Because experts will present both sides, the judge must be able to make an educated decision based on the underlying statistics. While this Article lays out exactly when each of the three criticisms is valid, it would likely take more than this brief exposure to guarantee that judges are prepared to make such an important decision. Judges must be educated on a variety of "scientific" topics to make any *Daubert* decision, including the admissibility of regression analyses. Many solutions to this education problem have been proposed. Scholars have called for judicial seminars to educate judges before litigation and independent research both before and during litigation.[248] Both of these methods could incorporate education on econometric

---

also FED. R. EVID. 702 (Testimony by Expert Witness).

247. *See* D.H. Kaye, *The Dynamics of* Daubert*: Methodology, Conclusions, and Fit in Statistical and Econometric Studies*, 87 VA. L. REV. 1933, 1985–87 (2001) (recognizing the importance of strictly analyzing statistics under *Daubert*).

248. *See* Edward K. Cheng, *Independent Judicial Research in the* Daubert *Age*, 56 DUKE L.J. 1263, 1270–75 (2007) (suggesting ways of "improving scientific admissibility decisions" through an educative approach). The use of court-appointed experts is discussed *infra* Part VI.B.

criticism. In particular, judicial conferences, such as the Science for Judges program,[249] which are already in place, could easily make this incorporation. This Article could also serve as a source for judges seeking independent research. However, because educating judges is often time-consuming and impractical,[250] we suggest that courts adopt a peer-review system. If a court adopts a peer-review system, such as the one proposed below, then the reliance on judicial education will be diminished.

B. Using Peer Review

Scholars concerned about the potential for junk science entering the courtroom through expert witness testimony have suggested several potential solutions to reduce those difficulties discussed in Part II.B. These proposed solutions include the use of court-appointed experts under Federal Rules of Evidence 706;[251] however, scholars have noted that this solution is not often practiced because it interferes with the adversary process.[252] Other solutions propose the establishment of a center of scientific experts that would act as a selection mechanism for potential court-appointed experts[253] and of an intermediary agency that answers blind technical questions for parties involved in litigation.[254] Lawrence Pinsky also suggested that expert testimony be peer reviewed in a more traditional sense.[255]

---

249. *See* Cheng, *supra* note 248, at 1273 (discussing judicial education programs).

250. *See id.* at 1273–74 (discussing difficulties with judicial education).

251. *See* FED. R. EVID. 706 (Court-Appointed Expert Witnesses). Many states also have a similar rule. *See* Cheng, *supra* note 248, at 1270 & n.21 (noting that many states permit court-appointed experts).

252. *See* Karen Butler Reisinger, Note, *Court-Appointed Expert Panels: A Comparison of Two Models*, 32 IND. L. REV. 225, 235–36 (1998) (discussing the opposition to court-appointed experts).

253. *See* Pinsky, *supra* note 53, at 545 (explaining solutions for assisting judges with handling complex scientific evidence).

254. *See* Christopher Tarver Robertson, *Blind Expertise*, 85 N.Y.U. L. REV. 174, 206–09 (2010) (detailing the concept of using an intermediary agency to "function[] as a broker between sponsors of research (e.g., plaintiffs) and potential expert witnesses (e.g., doctors)").

255. *See* Pinsky, *supra* note 53, at 558–62 (outlining the traditional methods of peer review).

In this solution, experts would present their reports during a pretrial hearing, and those reports would then be submitted to a committee for peer review.[256] The opinions of the committee would then be submitted to the judge and parties for review, and the judge would then make a decision about the admissibility of the evidence before trial.[257]

Pinsky's proposed solution of peer review is a viable solution that would assist judges in not only determining whether regression analyses should enter the courtroom under *Daubert* but also in determining whether the criticisms of regression analyses should be admitted. If this proposed peer-review process applied in an employment discrimination case with regression analyses, both the plaintiff's and the defendant's experts would present reports regarding the analysis they performed, including details on the variables included in the regression, the size of the sample, and how they calculated the statistical significance. In addition, the experts would also submit reports addressing their concerns with the opposing party's reports. Each of these reports would then be submitted for peer review. Economists skilled in regression analysis would undertake this peer-review process, submitting a response addressing actual deficiencies in the regression analysis and acknowledging whether the opposing expert's concerns have any merit. The judge would then take the peer-review commentary into account when determining not only whether the regression results should enter the courtroom but whether the opposing counsel and expert arguments that challenge the opposing party's regression should also be restricted. Alternatively, the court could simply rely on the peer-review commentary to expose the actual limitations of the regressions and not allow any additional criticisms to enter the courtroom.

Specifically, we propose a peer-review system in which both parties agree to provide a certain percentage of the fees the parties paid to their econometric experts to finance peer review.[258] Peer reviewers would be economists who do not

---

256. *See id.* at 543–44 (discussing a proposed peer-review solution to assist judges in determining the scientific validity of methodology employed by experts).

257. *See id.* (detailing the proposed peer-review solution).

258. If the plaintiff wins and the judge awards attorney's fees, then the

generally serve as litigation experts but who are experienced with peer review as academic scholars. Because academic economists serve as peer reviewers for academic journals for no compensation (occasionally a token payment is made), compensation on the order of 5% to 10% of the total billings by experts will provide adequate compensation to induce academic economists to participate on occasion. The judge would select the peer reviewers similar to how a judge chooses a court-appointed expert under Rule 706.[259] By selecting economists who do not generally serve as litigation experts, potential conflicts of interest will be avoided, as these economists will have no incentive to sway standards in expectation of benefiting from establishing statistical precedents.[260]

A summary of the proposed process follows: each party will submit one report; either both parties will submit their original analyses (if both plaintiff and defendant provide a primary analysis) or the plaintiff will submit a report and the defendant will submit its rebuttal report. At this point, both reports will be submitted to the peer reviewer who will advise the judge on the legitimacy of the reports and critiques. Based on the judge's assessment, invalid econometric critiques will be taken off the table, allowing parties to focus on only the appropriate and relevant issues in further rounds of expert reports and rebuttals and at trial.

judge could also award peer-review fees.

259. *See* FED. R. EVID. 706 ("The court may appoint any expert that the parties agree on and any of its own choosing.").

260. Anecdotally, many academic economists consider the litigation battle of experts to be difficult and often dishonest. As a result, economists who might be willing to be involved in litigation consulting if academic standards are maintained refuse to be involved as experts in anticipation that unscrupulous opposing experts (often professional consultants rather than academic economists) will launch erroneous and deceitful critiques. Because professional consultants have the advantage of greater litigation experience and are less concerned about their professional academic reputation, many qualified academic economists are driven out of the litigation arena. The proposed peer-review system would allow academic standards to enter courts' decision-making processes as an enhancement to the current adversarial process.

VII. Conclusion

Regression analysis has served an important role in employment discrimination cases for more than thirty-five years. Unfortunately, even though statistical evidence has become critical to the plaintiff's case in employment discrimination cases, regression analyses presented by a plaintiff to establish a prima facie case of disparate impact or disparate treatment do not increase the plaintiff's probability of prevailing in a case. Often the inability of valid regression analysis to assist a party is the result of the opposing expert's introduction of invalid econometric concerns. Because three of the most often cited econometric critiques are only valid in certain circumstances, judges must be aware that allowing such criticisms to enter the courtroom can influence the jury in a negative and unjustified way. As a result, judges should analyze the econometric criticisms presented under *Daubert* and limit the introduction of invalid econometric critiques. Because this solution likely requires extensive education of judges, courts should consider adopting a peer-review system that would rely on unbiased economists and guarantee that only valid regression results and valid econometric critiques enter the courtroom. Without such measures, flawed econometric critiques will continue to completely invalidate valid statistical evidence.

Appendix

Table A

Cases (listed from most recent to least recent)
Tabor v. Hilti, Inc., No. 09–cv–189–GKF–PJC, 2013 WL 4068781 (N.D. Okla. Aug. 12, 2013)
Parra v. Bashas', Inc., 291 F.R.D. 360 (D. Ariz. 2013)
Tabor v. Hilti, Inc., 703 F.3d 1206 (10th Cir. 2013)
Ellis v. Costco Wholesale Corp., 285 F.R.D. 492 (N.D. Cal. 2012)
Bolden v. Walsh Grp., 282 F.R.D. 169 (N.D. Ill. 2012)
Bazile v. City of Houston, 858 F. Supp. 2d 718 (S.D. Tex. 2012)
E.E.O.C. v. Bloomberg L.P., 778 F. Supp. 2d 458 (S.D.N.Y. 2011)
Wal-Mart Stores, Inc. v. Dukes, 131 S. Ct. 2541 (2011)
Powell v. Dall. Morning News L.P., 776 F. Supp. 2d 240 (N.D. Tex. 2011)
Rudwall v. Blackrock, Inc., No. C09-5176TEH, 2011 WL 767965 (N.D. Cal. Feb. 28, 2011)
Rollins v. Ala. Cmty. Coll. Sys., No. 2:09cv636-WHA, 2011 WL 3841954 (M.D. Ala. Aug. 26, 2011)
E.E.O.C. v. Bloomberg L.P., No. 07 Civ. 8383(LAP), 2010 WL 3466370 (S.D.N.Y. Aug. 31, 2010)
Aliotta v. Bair, 614 F.3d 556 (D.C. Cir. 2010)
Grant v. Metro. Gov't of Nashville, 727 F. Supp. 2d 677 (M.D. Tenn. 2010)
Bell v. Lockheed Martin Corp., 270 F.R.D. 186 (D.N.J. 2010)
Randall v. Rolls-Royce Corp., No. 1:06-cv-860-SEB-JMS, 2010 WL 987484 (S.D. Ind. Mar. 12, 2010)
E.E.O.C. v. Wal-Mart Stores, Inc., No. 6:01–CV–339–KKC, 2010 WL 583681 (E.D. Ky. Feb. 16, 2010)
Schanfield v. Sojitz Corp. of Am., 663 F. Supp. 2d 305 (S.D.N.Y. 2009)
Franklin v. Local 2 of the Sheet Metal Workers Int'l Ass'n, 565 F.3d 508 (8th Cir. 2009)
Ramirez v. Olympic Health Mgmt. Sys., Inc., 610 F. Supp. 2d 1266 (E.D. Wash. 2009)
Halabicky v. Univ. of Mich., No. 08–10065, 2009 WL 912629 (E.D. Mich. Mar. 30, 2009)
Boyd v. Interstate Brands Corp., 256 F.R.D. 340 (E.D.N.Y. 2009)
Puffer v. Allstate Ins. Co., 255 F.R.D. 450 (N.D. Ill. 2009)
Taylor v. United Parcel Serv., Inc., 554 F.3d 510 (5th Cir. 2008)
McClain v. Lufkin Indus., Inc., 519 F.3d 264 (5th Cir. 2008)
Taylor v. D.C. Water & Sewer Auth., 241 F.R.D. 33 (D.D.C. 2007)
Baylie v. Fed. Reserve Bank of Chi., 476 F.3d 522 (7th Cir. 2007)

Ellis v. Costco Wholesale Corp., 240 F.R.D. 627 (N.D. Cal. 2007)
Ram v. N.M. Dept. of Env't, No. CIV 05-1083 JB/WPL, 2006 WL 4079623 (D.N.M. Dec. 15, 2006)
Copeland v. CVS Pharm., Inc., No. CIVA 1:03CV3854 JOF., 2006 WL 2699045 (N.D. Ga. Sept. 15, 2006)
Wright v. Stern, 450 F. Supp. 2d 335 (S.D.N.Y. 2006)
E.E.O.C. v. Autozone, Inc., No. 00-2923 Ma/A, 2006 WL 2524093 (W.D. Tenn. Aug. 29, 2006)
Carpenter v. Boeing Co., 456 F.3d 1183 (10th Cir. 2006)
Nouri v. Boeing Co., 192 F. App'x 595 (9th Cir. 2006)
Colindres v. QuietFlex Mfg., 235 F.R.D. 347 (S.D. Tex. 2006)
Beck-Wilson v. Principi, 441 F.3d 353 (6th Cir. 2006)
Williams v. Boeing Co., No. C98-761P, 2006 WL 126440 (W.D. Wash. Jan. 17, 2006)
Jones v. GPU, Inc., 234 F.R.D. 82 (E.D. Pa. 2005)
Anderson v. Westinghouse Savannah River Co., 406 F.3d 248 (4th Cir. 2005)
Carlson v. C.H. Robinson Worldwide, Inc., No. Civ.02-3780 JNE/JGL, 2005 WL 758602 (D. Minn. Mar. 31, 2005)
Hnot v. Willis Grp. Holdings Ltd., 228 F.R.D. 476 (S.D.N.Y. 2005)
Obrey v. Johnson, 400 F.3d 691 (9th Cir. 2005)
McReynolds v. Sodexho Marriott Servs., Inc., 349 F. Supp. 2d 31 (D.D.C. 2004)
Cooper v. S. Co., 390 F.3d 695 (11th Cir. 2004)
Thomas v. Deloitte Consulting LP, No. 3–02–CV–0343–M, 2004 WL 1960097 (N.D. Tex. Sept. 2, 2004)
Morgan v. United Parcel Serv. of Am., Inc., 380 F.3d 459 (8th Cir. 2004)
Hershell Gill Consulting Eng'rs, Inc. v. Miami-Dade Cnty., 333 F. Supp. 2d 1305 (S.D. Fla. 2004)
E.E.O.C. v. Morgan Stanley & Co., 324 F. Supp. 2d 451 (S.D.N.Y. 2004)
Mehus v. Emporia State Univ., 222 F.R.D. 455 (D. Kan. 2004)
Dukes v. Wal-Mart Stores, Inc., 222 F.R.D. 137 (N.D. Cal. 2004)
United States v. City of Garland, No. Civ.A.3:98–CV–0307–L, 2004 WL 741295 (N.D. Tex. Mar. 31, 2004)
Anderson v. Boeing Co., 222 F.R.D. 521 (N.D. Okla. 2004)
United States v. Delaware, No. Civ.A. 01–020–KAJ, 2004 WL 609331 (D. Del. Mar. 22, 2004)
Cullen v. Ind. Univ. Bd. of Trs., 338 F.3d 693 (7th Cir. 2003)
Stone v. First Union Corp., 216 F.R.D. 540 (S.D. Fla. 2003)
Rhodes v. Cracker Barrel Old Country Store, Inc., 213 F.R.D. 619 (N.D. Ga. 2003)

Lanning v. Se. Pa. Transp. Auth., 308 F.3d 286 (3d Cir. 2002)
Hemmings v. Tidyman's Inc., 285 F.3d 1174 (9th Cir. 2002)
Webb v. Merck & Co., Inc., 206 F.R.D. 399 (E.D. Pa. 2002)
Meacham v. Knolls Atomic Power Lab., 185 F. Supp. 2d 193 (N.D.N.Y. 2002)
Karjala v. Winona State Univ., No. 99-1382(DWF/RLE), 2002 WL 27140 (D. Minn. Jan. 8, 2002)
Mengelkoch v. Bemidji State Univ., No. 99-1383(DWF/RLE), 2002 WL 27126 (D. Minn. Jan. 8, 2002)
Chalikia v. Moorhead State Univ., No. 99-2027(DWF/RLE), 2002 WL 27119 (D. Minn Jan. 8, 2002)
Cooper v. S. Co., 205 F.R.D. 596 (N.D. Ga. 2001)
Siler-Khodr v. Univ. of Tex. Health Sci. Ctr. San Antonio, 261 F.3d 542 (5th Cir. 2001)
Coleman v. Exxon Chem. Corp., 162 F. Supp. 2d 593 (S.D. Tex. 2001)
Reid v. Lockheed Martin Aeronautics Co., 205 F.R.D. 655 (N.D. Ga. 2001)
Fish v. St. Cloud State Univ., No. CIV 6:96-155DWF/RLE, 2001 WL 392038 (D. Minn. Mar. 26, 2001)
Fields v. Abbott Labs., No. 97C3882, 2001 WL 1663310 (N.D. Ill. Mar. 12, 2001)
Lavin-McEleney v. Marist Coll., 239 F.3d 476 (2d Cir. 2001)
Schanzer v. United Techs. Corp., 120 F. Supp. 2d 200 (D. Conn. 2000)
Guerrero v. Reno, No. 98 C 864, 2000 WL 1100400 (N.D. Ill. Aug. 7, 2000)
Morgan v. United Parcel Serv. of Am., Inc., 143 F. Supp. 2d 1143 (E.D. Mo. 2000)
Hayt v. New York Hosp., No. 97 CIV 5542 BSJ, 2000 WL 675347 (S.D.N.Y. May 23, 2000)
United States v. N.Y.C. Bd. of Educ., 85 F. Supp. 2d 130 (E.D.N.Y. 2000)
Derrickson v. Circuit City Stores, Inc., 84 F. Supp. 2d 679 (D. Md. 2000)
Muñoz v. Orr, 200 F.3d 291 (5th Cir. 2000)