Use of Statistics in Individual Disparate Treatment Cases

Yona Rozen and Adam Greenfield
Gillespie, Rozen & Watsky P.C.
3402 Oak Grove Ave. Suite 200
Dallas, Texas 75204
214-720-2009
The primary focus for this panel is use of statistics in class and pattern and practice cases, but it is interesting to also review the success and failures when individual plaintiffs seek to use statistics to prove their individual disparate treatment cases. With increasing regularity in recent years, statistical evidence and statistical experts are being used and implemented throughout the litigation process in an ever-expanding myriad of cases. The persuasiveness of statistical proof in employment discrimination matters is well established. See International Brotherhood of Teamsters v. United States, 431 U.S. 324 at 339-40 (1977) (stating that "our cases make it unmistakably clear that 'statistical analyses have served and will continue to serve an important role in cases in which the existence of discrimination is a disputed issue … We have repeatedly approved the use of statistical proof, where it reached proportions comparable to those in this case, to establish a prima facie case of racial discrimination in jury selection cases … Statistics are equally competent in proving employment discrimination). Also, see generally, Bazemore v. Friday, 478 U.S. 385, 404, 106 S.Ct. 3000, 92 L.Ed.2d 315 (1986).

The notion of using regression analysis in employment discrimination litigation dates at least to 1975 and the publication of a student note that advocated the idea. See, Note, Beyond the Prima Facie Case in Employment Discrimination Law: Statistical Proof and Rebuttal, 89 Harv. L. Rev. 387 (1975).

Plaintiffs historically have been able to use statistics in an effort to prove their claims of individual disparate treatment, but this proliferation of statistical proof has led to the battle over admissibility of such evidence being fought in the early stages of the case, through the trial phase. Pre-trial battles are typically fought during discovery, at the summary judgment stage, with Daubert challenges and Federal Rules of Evidence, Rule 403 objections. If statistical proof is successfully developed and introduced in those pre-trial stages, further issues arise as to how to effectively introduce and use such evidence at trial.

In this paper we attempt to look at some of the more recent cases,¹ where these battles have been waged and to draw practical suggestions as to how and when statistical proof can be utilized in individual disparate treatment cases with success, and/or to preview potential pitfalls.

In describing the persuasive power of numbers, particularly the use of statistics to bolster and/or weaken arguments, Benjamin Disraeli is frequently quoted as having said there are “lies,

¹ We do not purport to have included an exhaustive recitation of all cases dealing with statistics but have tried to include representative cases and to draw some suggestions and conclusions from cases in this area over the last several years.
damned lies, and statistics.” 2 Similarly one of the authors recalls a textbook used in her college statistics course entitled “How to Lie With Statistics”. 3 There is also the infamous line from computer science, “garbage in, garbage out” which perhaps can be updated to the current times as “don’t believe everything you read on the internet”. In our corner of the world, in litigation, we have the Federal Rules of Evidence and Civil Procedure, (or state rules) to act as peace keeper and police the battles over use of statistics.

In this paper we will examine issues regarding development of statistical evidence in discovery, use of that type of evidence to prove an individual disparate treatment case at the prima facie stage and to prove pretext (all at the summary judgment stage) and then when the case goes to trial, how effectively to present statistical evidence to the jury or court.

**Discovery**

In an individual disparate treatment case, building statistical proof can start with anecdotal evidence possessed by the plaintiff or provided to the plaintiff. 4 However, in all but the most unusual individual case, most of the information that would support statistical proof is going to reside in the hands of the defendant and will have to be dragged out in discovery. Such discovery battles are most frequently fought initially with written discovery. This scenario usually play out with the plaintiff requesting information or documents which are in turn objected to on the grounds of relevance, undue burden due to large geographic and temporal scopes, and raising confidentiality issues for other employees’ personal information (usually alleviated through protective orders and occasionally in camera inspection). 5

FRCP 26(b) clearly allows for “discovery of any matter relevant to the subject matter involved in the action.” The evidence is not necessarily admissible at trial, but it is discoverable

---

2 Checking for proper attribution of this quote led to numerous online discussions, showing that attribution, much like statistics, can be argued many ways. See for example: [http://www.york.ac.uk/depts/maths/histstat/lies.htm](http://www.york.ac.uk/depts/maths/histstat/lies.htm), site last visited, March 18, 2011 which suggests that it was likely not Disraeli who said this (although it has been attributed to him as far back as 1895).


4 Typically an employee may have been provided OWBPA information about others affected by a layoff or have access to institutional or legacy information regarding other similarly situated employees.

5 Duck v. Port Jefferson School Dist., 2008 WL 2079916, (EDNY 2008) (on motion to compel plaintiff was entitled to personnel records of other individuals who were similarly situated in that they retired and then were hired back to coaching positions or were hired in from “outside” the district for *in camera* inspection and the production if appropriate); Ladson v. Ulltra East Parking Corp., 164 F.R.D. 376 (S.D.N.Y., 1996), Mitchell v. Fishbein, 227 F.R.D. 239, 253 (S.D.N.Y.2005)
when relevant. While it is true that courts should prevent “general fishing expedition,” in the realm of individual disparate treatment cases, discovery is and should be liberal, and the courts have, and “may employ liberal civil discovery rules to obtain broad access to employer’s records.” This liberal approach to discovery has been specifically tailored to “statistical information concerning an employer’s general policy and practice concerning [protected class] employment,” as it may be relevant to a showing of pretext, “even in a case alleging an individual instance of discrimination rather that a ‘pattern and practice’ of discrimination.” It is of utmost importance to take this broad approach in discrimination cases because direct evidence of discrimination is seldom available and the statistical and pattern evidence needed to prove employment discrimination is often only available from the Defendant’s files.

In allowing or refusing to permit discovery on the grounds that the requested material is or is not relevant or unduly burdensome, courts are guided by what appears reasonable under the facts of the case. Under this ad-hoc approach, the courts force discovery requests to focus on the source of the complained discrimination; the employing unit, work unit, department, and sections in which there are employees who are situated similarly to the plaintiff.

In the absence of any evidence that there were hiring or firing practices and procedures applicable to all the employing units, discovery may be limited to plaintiff’s specific employing unit. However, discovery may be expanded beyond the plaintiff’s employing unit if the plaintiff can show the requested information is particularly cogent to the matter or if the plaintiff can show a more particularized need for, and the likely relevance of, broader information.

---

6 McDonnell Douglas Corp. v. Green, 411 U.S. 792, 805, 93 S.Ct. 1817, 36 L.Ed.2d 668 (1973); See also Sweat v. Miller Brewing Co., 708 F.2d 655, 658 (11th Cir.1983) (plaintiff, an older female labor relations representative at one brewing plant was entitled to reconsideration of motion to compel statistical discovery regarding labor relations representatives at other plants when circuit court reversed grant of summary judgment in view of disputed evidence of pretext because statistical evidence about employer’s policy and practice in minority employment could provide relevant evidence of pretext, even in an individual case); Rich v. Martin, 522 F.2d 33, 343-44 (10th Cir.1975); Haykel v. G.F.L. Furniture Leasing Co., 76 F.R.D. 386,391 (N.D.Ga. 1976)

7 Wards Cove Packing Co. v. Antonio, 490 U.S. 642, 657, 109 S.Ct. (1989); See also Weahkee v. Norton, 621 F.2d 1080, 1082 (10th Cir. 1980) (holding that in discrimination cases, “the discovery scope is extensive”).

8 Weahkee v. Norton, 621 F.2d 1080, 29 Fed. R. Serv. 2d 847 (10th Cir. 1980)

Discovery regarding an employer's facilities other than the one at which the plaintiff was employed may be denied on the ground that the information is not relevant.\(^{11}\) Such information is likely to be found irrelevant where a plaintiff has failed to show the existence of any particular employment practice and procedure applicable to all of the employing company's facilities and is challenging the employer's practices with respect to one facility only.\(^{12}\) Similarly, an employer will not be required to produce documents where the temporal scope of the request for the documents is overly broad.\(^{13}\)

**Prima Facie and Pretext Evidence**

The true battlegrounds on the use of statistics takes place when Plaintiffs use statistics to prove their prima facie case and pretext arguments. During this phase of proceedings, Defendants overwhelmingly object to Plaintiffs’ use of statistics particularly asserting the need of experts, further refinement of statistics, and the nature of the statistics in and of themselves.

At this point we are all familiar with *McDonnell Douglas* shifting burden:

1. Plaintiff has a burden in establishing a prima facie case raising an inference of discrimination;

2. The burden then arises for Defendant to articulate, through admissible evidence, a legitimate and non-discriminatory reason for its action

3. Plaintiff must then prove that the employer’s articulated non-discriminatory motive is a mere pretext for discrimination.

Statistical proof can be offered at the prima facie stage in an individual disparate treatment case. In individual disparate treatment cases, where the employer's motivation is the


ultimate issue, the plaintiff's prima facie case eliminates only the most common nondiscriminatory reasons for differences in treatment. It is not until the employer responds by articulating the “real” reason for its different treatment of plaintiff that the plaintiff ultimately proves discrimination by eliminating the alleged non-discriminatory reason, as well.

A plaintiff may use statistical evidence to establish a prima facie case of discrimination or subsequently to prove pretext. Statistics may be used to establish an individual plaintiff's claim that a pattern or practice of discrimination existed, or simply to bolster the plaintiff's circumstantial evidence of individual disparate treatment. Evidence establishing a discriminatory pattern is probative of motive and can therefore create an inference of discriminatory intent with respect to the individual employment decision at issue\(^\text{14}\). However, this statistical evidence, in order to be relevant, must meet certain criteria in its source and refinement.

In relation to the source of statistics, Plaintiffs seeking to rely on statistical evidence must be similarly situated to those individuals in the statistical group studied, so as to render the analysis relevant\(^\text{15}\). In relation to refinement, raw statistics are rarely probative of anything. To be meaningful, further refinement and analytical analysis (if not always expert analysis) is required\(^\text{16}\). However, Districts vary on the need of experts and the refinement of statistics when making out a prima facie case\(^\text{17}\). These different approaches are cited by the district court in its opinion in *Nicholls*:

There is room for debate about whether such raw statistical data ever can make out a prima facie case. *Compare Smith v. Xerox Corp.*, 196 F.3d 358, 370 (2d Cir.1999) (noting that in disparate treatment cases, “a plaintiff may [ ] present statistical findings as circumstantial evidence of intentional discrimination”), *overruled on other grounds by Meacham v. Knolls Atomic Power Lab.*, 461 F.3d 134 (2d Cir.2006), *vacated*, 554 U.S. 84, 128 S.Ct. 2395, 171 L.Ed.2d 283 (2008); *Stratton v. Dep't for the Aging for N.Y.*, 132 F.3d 869, 877 (2d Cir.1997)


(noting that expert statistical analysis is not required in individual disparate treatment cases); 
*Bruno v. W.B. Saunders Co.*, 882 F.2d 760 (3d Cir.1989) (noting that statistical evidence in individual disparate treatment cases “need not be ... finely tuned” and that “the usefulness of statistics will depend primarily upon their relevance to the specific decision affecting the individual plaintiff” (internal quotation marks omitted)), with *Miller v. Nat'l Ass'n of Sec. Dealers, Inc.*, 703 F.Supp.2d 230, 245 (E.D.N.Y.2010) (“Statistical information ... may be considered as circumstantial evidence supporting the necessary inference of discrimination. However, statistical proof alone cannot ordinarily establish a prima facie case of disparate treatment.” (internal quotation marks, citations, and alteration omitted)); *Baron v. N.Y.C Dep't of Educ.*, No. 06-CV-2816, 2009 WL 1938975, at *6 (E.D.N.Y. July 7, 2009) (noting that while statistical evidence, when combined with other evidence of discriminatory animus, may be sufficient to establish a prima facie case, “statistics alone are insufficient in a disparate-treatment [case] because an individual plaintiff must prove that he or she in particular has been discriminated against” (internal quotation marks and alteration omitted) (emphasis in original)).

2011 WL 180565 at *8

The debate among various courts as to whether or not an expert witness is necessary to present statistical evidence, brings out a major issue with use of statistics in an individual disparate treatment case, and that is cost. Unlike a class or pattern and practice case, it is rare that an individual plaintiff would have the resources to allow retention of an expert and unless the plaintiff is a high earner, an individual disparate treatment case may not involve potential damages of a sufficient magnitude to support or justify such an expense. This is probably why many of the cases involve the plaintiff or their counsel, trying to make sufficiently sophisticated statistical presentations to withstand objections by defendants or dismissive treatment by judges.

With respect to model specifications (all relevant explanatory variables), it is appropriate to talk about *Daubert*, which requires that regression analysis must meet the standards that economists would apply to their non-litigation research. Omitted variables can be fatal to the ability of statistical analysis to inform the jury and illustrates what “undesirable results” that occur when regression models are incorrectly specified. When an important explanatory variable is omitted, the statistical analysis becomes unreliable. It not only misleads, it lacks the capacity to inform, so it cannot be shown to have probative value. In Rule 403 terms, it has no probative value, but it surely has the capacity to misinform the jury, so the latter danger must substantially outweigh the nonexistent former probative value.\(^\text{18}\). However, this is in contrast to *Abrams v.*

---

Lightolier Inc., 50 F.3d 1204, 1217 (3d Cir. 1995) (citing Bruno v. W.B. Saunders Co., 882 F.2d 760, 766 (3d Cir. 1990) which notes that statistics in an individual disparate treatment case, in contrast to a class-action or pattern-and-practice case, “need not be so finely tuned”. In addition to missed variables, defendants also heavily object to the data in and of itself being misleading. This objection comes into play when the plaintiff uses a sample size that is deemed too small and thus misleading. This often raised issue of sample size is problematic in terms of using statistical evidence in an individual disparate treatment case since more often than not, there are a fairly small number of similarly situated individuals.

Infrequently, statistical evidence alone has been used to present a prima facie case for an individual claim. These statistics must pertain to the decision being challenged, and must be sufficient to raise the inference of disparate treatment discrimination with regard to the plaintiff’s situation. However, statistical proof is usually used as adjunct evidence in individual disparate treatment cases. Normally, a statistical rebuttal to a prima facie case in an individual disparate treatment claim will not be effective unless the individual claim is the rare one that relies solely on statistics. While the Supreme Court found statistical data to be not wholly irrelevant in addressing an individual claim of disparate treatment, and declined to say that such proof would have absolutely no probative value, the court concluded that a balanced workforce cannot immunize an employer from liability for specific acts of discrimination.

The purpose of the order of proof in individual disparate treatment cases is to narrow the issue of the employer's intent gradually. Thus, the ultimate issue of discriminatory motive is most often decided, assuming plaintiff made a prima facie showing, at the final stage, when the plaintiff must prove that the employer's articulated non-discriminatory motive is a mere pretext for discrimination. The use of statistical evidence to show pretext had earlier been approved in McDonnell Douglas Corporation v. Green, 411 U.S. 792, 804-05, 93 S.Ct. 1817, 1825-26

---

19 Conlay v. Baylor College of Medicine, 688 F.Supp.2d 586, (S.D. Tex. 2010)(sample of 7 people too small); Shutt v. Sandoz Crop Protection Corp., 944 F.2d 1431, 1433-34 (9th Cir.1991) (sample of 11 is too small to establish a statistical pattern of age discrimination); Jackson v. Univ. of New Haven, 228 F.Supp.2d 156, 164-65 (D.Conn.2002) (noting, in disparate impact case involving sample size of 14 people, that “exceedingly small sample sizes often result in statistically unreliable evidence”); Waris v. HCR Manor Care, (sample size of five is too small). 20 Davis v. Califano, 613 F.2d 957 (D.C. Cir. 1979)
(1973), where the Court stated that “statistics as to petitioner's employment policy and practice may be helpful to a determination of whether petitioner's refusal to rehire respondent in this case conformed to a general pattern of discrimination.”

A plaintiff may use statistical evidence to establish pretext, but most cases fail when all the plaintiff has is statistical evidence without more. Under the McDonnell Douglas individual disparate treatment model, statistical evidence is “only one small part of a substantial web of evidence indicating pretext”. However, statistics in an individual disparate treatment case, in contrast to a class-action or pattern-and practice case, “need not be so finely tuned”, and can play a large role as circumstantial evidence in proving pretext.

Most cases where statistical evidence is all that plaintiff offers fail and are dismissed on summary judgment if they don’t present additional evidence. Those cases using statistics that have been successful, present the statistical evidence, together with other evidence to show that the defendant’s asserted legitimate, non-discriminatory reasons are pretext.

The Supreme Court’s decision in Sprint/United Management Co. v. Mendelsohn, 552 U.S. 379, 128 S.Ct. 1140 (U.S. 2008), may provide the basis for additional support for use of statistical evidence in individual disparate treatment cases. In Mendelson the Court granted certiori and determined whether, in an employment discrimination action, the Federal Rules of Evidence require admission of testimony by nonparties alleging discrimination at the hands of persons who played no role in the adverse employment decision challenged by the plaintiff [so called “me-too evidence”]. The Court found that “the question whether evidence of discrimination by other supervisors is relevant . . . is fact based and depends on many factors, including how closely related the evidence is to the plaintiff’s circumstances and theory of the case. Applying Rule 403 to determine if evidence is prejudicial also requires a fact-intensive,

24 Boone v. Clinton, 675 F.Supp.2d 137, (D.D.C. 2009) -- While the analysis here might be different if this statistical evidence were all that Boone presented to survive the Department's motion for summary judgment, see Krodol, 748 F.2d at 710, Boone presents additional evidence from which a trier of fact could reasonably conclude that race motivated her non-selection; Jones v. Mukasey, 565 F.Supp.2d 68, 78 (D.D.C.2008) -- denying summary judgment because evidence of a statistically significant deviation between the hiring of African-American and white applicants, in conjunction with other evidence, raised an issue of pretext; Nicholls v. Philips Semiconductor Mfg.; Conlay v. Baylor College of Medicine; Robertson v. Dodaro, Foster v. Mid State Land & Timber Co., Inc. – Did not provide additional evidence and did NOT survive MSJ.)
context-specific inquiry.” This was so even though none of the five witnesses worked in the same group with Mendelsohn, nor had any of them worked under or reported hearing discriminatory remarks the supervisors in her chain of command. Following Mendelsohn, one interesting approach one court has taken can be found in Miller v. Love's Travel Stops & Country Stores, Inc., 2008 WL 2079961, (W.D.Okla. 2008), which may help a plaintiff use statistical evidence to get past summary judgment in an individual disparate treatment claim.

In Miller, Defendants sought to exclude evidence and argument regarding employment discrimination claims and administrative charges filed by other former employees who were not similarly situated to plaintiff and had no personal knowledge of the circumstances surrounding plaintiff’s termination. Plaintiff opposed the motion by arguing that evidence of defendant's age-based terminations of other employees was relevant, and necessary, to prove plaintiff's allegation that defendant engaged in a company-wide discriminatory practice of terminating older workers, including plaintiff.

In denying Defendant’s First Motion In Limine, the court held that they had “insufficient information to conclude that other employees' terminations were logically or reasonably tied to [Defendant’s] decision to terminate Plaintiff,” and “In addition, performing a Rule 403 balancing of probative force against prejudicial effect will require the Court to evaluate the proposed testimony within the context of other trial evidence, which remains largely unknown to the Court.”

The Court then directed Plaintiff to “inform the Court during the trial before he calls a former employee to testify about the witness's alleged discriminatory termination by Defendant. At that time, the Court will conduct an inquiry outside the presence of the jury into facts necessary to determine the admissibility of the witness's anticipated testimony. If necessary, this inquiry may include a preliminary examination of the witness pursuant to Fed.R.Evid. 104. The Court will expect Plaintiff to make a particularized showing as to the relevance of the witness's proposed testimony, and then will permit the parties to present their arguments regarding Rule 403 considerations. If either party intends to request a limiting instruction to mitigate any
prejudice to Defendant of admitting such evidence, a proposed jury instruction should be submitted in writing for consideration by the Court in advance of the presentation of evidence.” 2008 WL 2079961 at *2.

The language in Mendelsohn, which requires a fact-based review of a number of factors (the circumstances and theory of the plaintiff’s case) can be used to support an argument for similar fact-based review of evidence which would support statistical evidence of discrimination beyond the plaintiff’s individual circumstances in an disparate treatment case. The court’s method for handling these issues set forth above in Miller v. Love's Travel Stops & Country Stores, Inc. indicates that procedures and processes can be fashioned to balance the competing interests in connection with this type of evidence. It does not have to wait until the time or trial as is demonstrated in some of the processes crafted in the discovery cases discussed above where the court required protective orders, in camera inspection and such. Also, something to be considered is whether the statistical evidence needs to have identifying facts associated with it or whether discovery of records reflecting employee A, B, C etc. is sufficient for the purposes of establishing the statistical evidence.

Where the plaintiff survives summary judgment and makes it to trial in a disparate treatment case, how should statistical evidence be presented to a jury? As noted above, in many cases a statistical expert is beyond the plaintiff’s means. In those circumstances the precedent suggesting that statistical evidence in a disparate treatment case “need not be so finely tuned” is helpful. It is important though to have thought about how to present the statistical evidence from the beginning so that sufficient details were developed in discovery to flesh out the details and put the statistics in context. It is important to introduce the underlying evidence that supports the statistical evidence so it is in the record but in terms of presenting it to the jury graphics and charts are extremely helpful in bringing the statistical evidence to life. Use of summaries are critical so that all the evidence which may be drawn from various sources is in one place where it can be highlighted for the jury and where they can find it. Also the lessons drawn from the difference in success between a case presented solely based on statistical evidence as compared to the cases where statistical evidence was presented in addition to various other evidence of pretext, should not be forgotten. It is probably best to lay out a lot of the other evidence and then have the statistical evidence as the “icing on the cake”.

743
In sum, an individual disparate treatment case where the only evidence of discrimination is based on statistical proof is unlikely to be successful but strong statistical evidence, can be appropriate and persuasive in an individual case, when it is presented in addition to other substantial evidence of pretext and has been developed in substantial detail to add additional support to plaintiff’s case.