

## COMMERCIAL DIVISION UPDATE

## Expert Analysis

# The Use of Statistical Sampling as Evidence

In a case involving failed mortgage-backed securities, the New York County Commercial Division has recently ruled that expert testimony of statistical sampling will be admissible at trial.<sup>1</sup> In *MBIA Ins. Corp. v. Countrywide Home Loans Inc.*, Justice Eileen Bransten granted plaintiff's motion in limine to allow such evidence after finding that the proffered methodology was scientifically accepted and reliable, and that a sampling of the mortgages at issue would be appropriate as the case involved more than 368,000 loans.

While recognizing that defendants' concerns about the sampling methodology were not without merit, the court noted that it was not necessarily endorsing the plaintiff's method as better than any other method of proof. The court observed that, ultimately, the plaintiff must show at trial that the sample chosen is actually statistically significant and must "convince the trier of fact, whether it be a jury or this court, that it has proven each element of its separate claims.... Sampling itself is not proof, but merely a vehicle to present evidence."<sup>2</sup>

### History in New York Courts

Statistical sampling is a scientific methodology by which one draws conclusions about a large population of data by measuring and analyzing a smaller, representative sample of the population. When the sample is randomly selected and of sufficient size so as to achieve statistical significance, statisticians may confidently make inferences about the larger population by reviewing the sample. As such, statistical sampling can provide an efficient way to estimate accurately larger populations of data, and has been utilized across many spectrums outside of the courtroom, including election polling, television ratings, unemployment surveys and analyses of public health issues.

In the courtroom, the applicability of inferential statistics to present evidence from a large population of data has long been accepted, provided that the party seeking to use statistical sampling can demonstrate that the proposed sample is truly representative of the larger population and that the methodology used by the expert is valid and reliable.<sup>3</sup>

In 1856, for example, the New York State Court of Appeals acknowledged the potential utility of statistical sampling evidence in a breach of warranty case.<sup>4</sup> In *Muller v. Eno*, the New York Court of Appeals



By  
**George  
Bundy Smith**



And  
**Thomas J.  
Hall**

reversed the trial court which had instructed the jury that it could not infer that all of the cotton cloth bales that plaintiff bought from defendants were damaged based on the plaintiff's testimony that he inspected only several of the bales and found such damage.

In so ruling, the Court held that the plaintiff was not bound to produce the best evidence, and that there "was no rule of law [that] would be violated by leaving [the jury] to infer, if they pleased to do so, that all the [bales] purchased...were unsound to the same extent as those particularly examined. [We] do not say that the inference would be a very strong one, or that it might not be overcome by what was considered more direct evidence on the other side; but [we] think the whole question was for the jury."<sup>5</sup> Since *Muller*, New York state and federal courts have repeatedly permitted the use of statistical sampling as evidence in cases involving significant amounts of data.<sup>6</sup>

Statistical sampling can provide an efficient way to estimate accurately larger populations of data.

As New York case law has evolved, three factors have come to dominate the admissibility analysis. First, the court must find that the expert testimony relying on statistical sampling is based on scientific principles or procedures that have been "sufficiently established to have gained general acceptance in the particular field to which it belongs."<sup>7</sup> Second, the court must examine the "specific reliability of the [actual] procedures followed [by the expert] to generate the evidence proffered," by asking "whether the accepted methods were appropriately employed in a particular case."<sup>8</sup> And third, the court must determine that statistical sampling evidence would be appropriate for the specific issue for which it is to be used.<sup>9</sup>

### 'MBIA Ins. v. Countrywide'

In *MBIA Ins. Corp.*, plaintiff MBIA Insurance asserted claims for breach of contract and fraud against various Countrywide entities after MBIA

allegedly suffered losses on insurance it had issued on 15 securitized pools of residential mortgage loans. Plaintiff alleged that defendants made materially false and misleading representations that induced plaintiff to provide that insurance, and also that they breached warranties applicable to the 368,000 mortgage loans in the 15 securitized pools.

Plaintiff sought to use statistical sampling to "streamline the presentation of evidence at trial, conserve the time and resources of th[e] Court and the parties, and result in a speedy and efficient disposition of the matters at issue without prejudicing either party."<sup>10</sup> According to the plaintiff, sampling is particularly appropriate for this case as there are over "368,000 loans and corresponding loan files, each containing a substantial amount of origination, underwriting, and servicing information [and] the alternative [of] present[ing] evidence on a loan-by-loan basis...is virtually impossible, and would unnecessarily prolong the trial, delay the resolution of [the] claims, and result in an enormous drain on the Court's time and resources."<sup>11</sup>

In permitting the use of statistical sampling, Justice Bransten reviewed and rejected each of defendants' arguments against admissibility. Defendants argued, for example, that no prior cases had permitted statistical sampling evidence involving mortgage-backed securities or in a closely analogous context.<sup>12</sup> Defendants also contended that since plaintiff's claims raised a wide variety of factual circumstances, sampling would be an inappropriate method to prove the claims.<sup>13</sup> Plaintiff countered that the accepted techniques of statistical sampling are based on "mathematical principles and are the same regardless of whether the population being sampled is widgets, tires, or loans," and thus its admissibility can be supported by other types of cases.<sup>14</sup>

The court found that the admission of scientific evidence is not limited to certain cases and, considering the fact that mortgage-backed securities are a relatively new device, noted that it was also not surprising that plaintiff did not cite a case allowing such evidence in this context. In each case, the court found, scientific evidence is to be reviewed of its own accord.<sup>15</sup>

Justice Bransten then concluded that the proposed statistical sampling is scientifically acceptable and generally admissible under the *Frye*<sup>16</sup> standard. The court cited to cases in a variety of contexts, finding that statistical sampling is a widely used method to present evidence from a large population of data and is not a novel concept. According to the court, it is undisputed that statistical sampling is generally accepted in the scientific community, that "[s]cientific literature and testing is replete with the use of statistical sampling," and that "the validity of properly conducted sampling is not a question for debate."

GEORGE BUNDY SMITH and THOMAS J. HALL are partners with Chadbourne & Parke. Mr. Smith served as an Associate Judge on the New York Court of Appeals from 1992 to 2006. CHRISTOPHER CUSMANO, a litigation associate at the firm, assisted with the preparation of this article.

In fact, the court noted that defendants objected to neither the novelty nor the general acceptance in the scientific field of statistical sampling.<sup>17</sup>

Finally, the court determined that plaintiff's expert's proposed methodology adhered to the accepted methods and would be reliable. Plaintiff's expert proposed to undergo a sampling exercise twice. In each such exercise, the expert would randomly choose 400 loans from each of the 15 securitizations at issue. According to plaintiff's expert, the resulting sample size of 6,000 loans from the larger population of 368,000 loans would yield a confidence level of 95 percent and a margin of error of 5 percent.<sup>18</sup>

After generating the random samples, the expert would then stratify the samples into smaller subgroups based on a variety of factors including the particular claim to which the loan was related and other variables considered by defendants when making lending decisions, such as the borrower's credit score. This process is designed to ensure representativeness and reduce the variability of estimates, making the results more precise.

Finally, after stratification, the expert would determine: (i) the total number of loans in the population that fall into each of the strata; (ii) the total number of loans in the sample that fall into each of the strata; (iii) the proportion of loans in the population from each stratum; and (iv) the breach rate across strata.<sup>19</sup> The expert would then testify with a 95 percent confidence level that this sample breach rate across all strata is a proper estimate of the breach rate in the entire population of the 368,000 loans.

Defendants vigorously opposed the plaintiff's expert's proposed methodology as flawed since the "proposal fail[ed] to use appropriate statistical methods to extrapolate liability for Plaintiff's claims and any damages associated with those claims."<sup>20</sup> More specifically, defendants argued that plaintiff's expert failed to demonstrate that the sample was of a sufficient size so as to be statistically reliable, since the proposed methodology is based on "an incorrect assumption that the assessment of the sampled loans will involve only discrete 'yes' or 'no' outcomes," and does not take into account whether a particular loan may be compliant with the warranties in certain respects but not others.<sup>21</sup>

In addition, defendants stated that because each of the 15 securitizations included loans that were "heterogeneous and var[ie]d widely based upon at least ten different characteristics" and the stratification process failed to take all of these factors into account, the samples would be unrepresentative and unreliable.<sup>22</sup>

Ultimately, Justice Bransten found that the proposed methodology was admissible as plaintiff demonstrated it was scientifically accepted, valid and reliable. The court was careful to note, however, that it was making no finding that plaintiff's proposed method is without flaw or insusceptible to challenge. While acknowledging that defendants raised significant challenges to plaintiff's methodology, the court found that they went to the weight, not the admissibility, of the evidence.

In fact, the court suggested that "[a]s Plaintiff may present its case as it chooses, so may Defendants rebut Plaintiff's proffered arguments through Defendants own sampling chosen in a statistically valid manner [and t]he ultimate trier of fact will then decide the issue."<sup>23</sup> In the end, the court found that statistical sampling would be a reliable, economical and timely alternative to individualized litigation as to each of the 368,000 loans in question.

## Not Always Appropriate

Although statistical sampling is "generally accepted" in the scientific community and often found admissible by courts, there may still be instances where it will not be deemed "appropriate" for a particular commercial case. In *Alix v. Wal-Mart Stores Inc.*,<sup>24</sup> for example, Justice Richard M. Platkin of the Albany County Commercial Division determined that plaintiffs' proffered use of expert testimony based on statistical sampling was not an appropriate substitute for individualized proof in a case based on alleged violations of the New York Labor Law statute.

Plaintiffs sought to establish a class action for all Wal-Mart employees, alleging that defendant systematically deprived hourly workers of wages through a variety of unlawful practices. Such practices allegedly included requiring employees to work "off-the-clock" during breaks, and before or after shifts, and also falsifying time records by using "one-minute punch-outs," which were allegedly fictitious time entries entered by managers to reflect a punch-out one minute after an employee's shift had started if the employee had failed to properly swipe out after their shift had ended, thereby causing the employee to fail to be paid for the entirety of the shift.<sup>25</sup>

Justice Bransten found that the proposed methodology was admissible as plaintiff demonstrated it was scientifically accepted, valid and reliable. The court was careful to note, however, that it was making no finding that plaintiff's proposed method is without flaw or insusceptible to challenge.

Justice Platkin rejected plaintiffs' attempt to use statistical sampling evidence to prove the alleged violations, instead requiring each employee to prove his or her claims separately. The court explained that to establish a prima facie case of violations under the Labor Law statute, plaintiffs needed to adduce specific evidence as to which associates worked off the clock, on what occasions, and for how long. The court found, however, that statistical sampling would not be appropriate to capture this individualized proof. For example, the court noted that the facts and circumstances surrounding the allegedly unpaid work vary substantially from associate to associate.

The court cited, for example, the vast disparities in the nature and factual circumstances surrounding the claims of the two named plaintiffs, and found that the extrapolation of evidence for all employees from a random sample would be problematic, and not appropriate in the case. Moreover, the court expressed concern that, even if statistical sampling could identify irregularities in time-recording, under the statute, liability could only be imposed through an individualized account for each employee raising a claim, since the fact-finder must eventually determine exactly how much of unrecorded time was spent on employment matters for which compensation was deserved.<sup>26</sup>

The court also commented on the plaintiffs' reliance on federal and secondary authority supporting the use

of statistical sampling as evidence of individualized proof in similar cases. According to the court, such reliance rests on shaky evidentiary ground because New York state courts have not adopted the more flexible admissibility rules found in federal courts and some other states, instead relying on the *Frye* standard. Finally, the court mentioned considerations of fundamental fairness, as defendant sought an opportunity to offer numerous individualized, employee-specific defenses as to liability and damages, but would be unable to do so if the plaintiffs relied on statistical sampling evidence as a substitute for individualized proof.<sup>27</sup>

## Conclusion

The use of statistical sampling can significantly ease the proof required of a plaintiff at trial. As demonstrated by *MBIA Ins. Corp.* and *Alix*, the relative appropriateness of statistical sampling to the issues in a particular case may be the most determinative factor as to whether sampling evidence is admissible. Litigation counsel must be careful to use a methodology appropriate to the circumstances of the case as was done in *MBIA Ins. Corp.* Even where statistical sampling evidence is admitted, however, the opposing party will be entitled at trial to challenge the reliability of that evidence so as to undermine the weight that may be given to it by the finder of fact.

1. *MBIA Ins. Corp. v. Countrywide Home Loans Inc.*, NYLJ, Jan. 14, 2011, No. 602825/08, 2010 WL 5186702 (N.Y. Co. Dec. 22, 2010) (Bransten, J.).

2. *Id.* at 6.

3. *Id.* at 4; McCormick on Evidence §209 (6th ed. 2006) ("[T]he battle is not usually over the admissibility of statistical evidence in general... Rather, the battle lines are drawn where it comes to the admissibility of obviously flawed applications, to the weight that should be given the evidence, and to transforming methods and conventions of statistical inference into rules of law"); *In re Mercy Hosp. of Watertown*, 79 N.Y.2d 197, 203-04, 581 N.Y.S.2d 628, 631-32 (1992).

4. *Muller v. Eno*, 14 N.Y. 597 (1856).

5. *Id.* at 600, 603-04.

6. *In re Mercy Hosp. of Watertown*, 79 N.Y.2d at 203-04, 581 N.Y.S.2d at 631-32; *Sunset Taxi Co. Inc. v. Blum*, 73 A.D.2d 691, 692, 423 N.Y.S.2d 231, 232 (2d Dept. 1979) ("[T]he contention that sample evidence provides no evidence for the goods or claims not sampled was rejected long ago"); *In re World Trade Ctr. Disaster Site Litig.*, 598 F.Supp.2d 498, 503-05 (S.D.N.Y. 2009) (Case Management Order that included a "methodology for sampling in relation to the general run of cases, severe, mild, and everything between, in order that rulings on liability, damages, and responsibility might be extended from the particular case in which rulings are made to the rest of the cases" was a "flexible, fair, and efficient plan to move [9,090] cases through discovery and to trial in reasonable time"); *Rosado v. Wyman*, 322 F.Supp. 1173, 1180 (E.D.N.Y. 1970) ("Sampling has long been considered an acceptable method of determining the characteristics of a large universe... Such mathematical and statistical methods are well recognized by the courts as reliable and acceptable in determining adjudicative facts"), *aff'd*, 437 F.2d 619 (2d Cir. 1970).

7. *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

8. *Parker v. Mobil Oil Corp.*, 7 N.Y.3d 434, 447, 824 N.Y.S.2d 584, 589 (2006) (citing *People v. Wesley*, 83 N.Y.2d 417, 429, 611 N.Y.S.2d 97, 103-04) (1994)).

9. *MBIA Ins. Corp.*, 2010 WL 5186702, at 3.

10. Plaintiff's Mem. of Law dated April 30, 2010 at 1.

11. *Id.* at 1-2.

12. Defendants' Supp. Mem. of Law dated Aug. 25, 2010 at 12.

13. *Id.* at 13.

14. Plaintiff's Supp. Mem. of Law dated Sept. 2, 2010 at 4.

15. *MBIA Ins. Corp.*, 2010 WL 5186702, at 3.

16. *Frye*, 293 F. at 1014.

17. *MBIA Ins. Corp.*, 2010 WL 5186702, at 4-5.

18. *Id.*

19. Plaintiff's Supp. Mem. of Law dated Sept. 2, 2010 at 10.

20. Defendants' Supp. Mem. of Law dated Aug. 25, 2010 at 15.

21. *Id.*

22. *Id.* at 16.

23. *MBIA Ins. Corp.*, 2010 WL 5186702, at 6.

24. *Alix v. Wal-Mart Stores Inc.*, 16 Misc. 3d 844, 838 N.Y.S.2d 885 (Albany Co. 2007) (Platkin, J.).

25. *Id.* at 845-46, 838 N.Y.S.2d at 887-88.

26. *Id.* at 854-57, 838 N.Y.S.2d at 893-96.

27. *Id.* at 856-57, 838 N.Y.S.2d at 895-96.