



ELSEVIER

Journal of Econometrics 113 (2003) 69–82

JOURNAL OF  
Econometrics

www.elsevier.com/locate/econbase

# Issues arising in using samples as evidence in trademark cases<sup>☆</sup>

Joseph L. Gastwirth<sup>a,b,\*</sup>

<sup>a</sup>*Department of Statistics, George Washington University, Fungler Hall 315, Washington, DC 20052, USA*

<sup>b</sup>*Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892, USA*

Accepted 16 July 2002

---

## Abstract

This article reviews the use of survey evidence in cases involving alleged infringement of trademarks. Because courts do not admit “hearsay” evidence, at first, courts were reluctant to accept evidence based on a sample or portion of the relevant population. Some early cases that led to the legal acceptance of samples and surveys are described. The role of survey evidence in a variety of types of trademark litigation, e.g., dilution or deceptive advertising are illustrated with data from actual cases.

© 2002 Elsevier Science B.V. All rights reserved.

*JEL classification:* C4; K41

*Keywords:* Law and economics; Sample evidence; Sensitivity analysis; Survey evidence; Trademark litigation

---

## 1. Introduction

Many of the most important data sets by economists are based on statistical samples. For example, the labor-force data and related income distribution and poverty count are obtained from a monthly sample of households and the Consumer Price index is based on a sample of prices of items bought by the “typical” household. The legal system was slow to accept information based on samples but more recently survey data has been

---

<sup>☆</sup> Research supported in part by a National Science Foundation Grant.

\* Corresponding author. Department of Statistics, George Washington University, Fungler Hall 315, 20052 Washington, DC 20052, USA.

*E-mail address:* jlgast@gwu.edu (J.L. Gastwirth).

used in a variety of cases. For example, samples of financial records are used to estimate hospital overcharges in Medicaid or Medicare cases (Heiner et al., 1984), economic damages in equal employment and lost profits cases (Walker and Monahan, 1998; Hall and Lazear, 2000) and estimating the tax basis of stock acquisitions. Samples of the assessed value and sales prices of farms or homes are utilized in cases concerning the fairness of property assessments (Gastwirth, 1988, 1992). Samples are also used in criminal cases, e.g., estimating drug quantities (Aitken, 2000; Izenman, 2000). Two areas of special importance to economics are the use of surveys to define the relevant market area and the characteristics that affect consumer preferences in anti-trust cases (King, 1986; Diamond, 1994, 2000) and to assess whether a product infringes upon an established trademark (Gastwirth, 1988; Diamond, 1994, 2000). After reviewing early cases involving the acceptance of samples by courts, this article describes the evidentiary role of surveys in Lanham Act litigation.

Some early cases are described in Section 2 to illustrate how it took some time for the legal process to accept sample information in lieu of data for the entire population. Section 3 is devoted to the different uses of surveys in trademark and misleading advertising cases concerning possible violations of the Lanham Act. This is the legal area that most frequently relies on survey evidence. The importance of measuring the appropriate concept as well as properly designing the data collection process so the sample represents the relevant population of potential consumers data is emphasized.

## 2. The gradual acceptance of samples by the US legal system

The Federal Rules of Evidence, especially Rule 703, which focuses on whether the facts or data submitted are “of a type reasonably relied on by experts in the particular field in forming opinions or inferences”, allows survey and other sample-based evidence to be used by experts in forming their opinion. The rule allows experts to rely on information that is inadmissible itself in forming their opinions. Before these rules were adopted, some courts were reluctant to consider evidence based on samples. The *Sears Roebuck vs. City of Inglewood* case, discussed by Sprowls (1956–57) and Gastwirth (1988, p. 495), concerned a tax refund request made by Sears. The city imposed a tax of 1/2 of 1% on all sales made by stores within the city limits but excluded sales made to persons not living in the city. During a routine audit the company found it had overpaid its tax because it relied on an erroneous definition of an “out of city” sale. To support its request for a refund of \$27,000 Sears relied on a random sample of sales slips. The sampling was done by choosing 3 days from each quarter of a year during the time period. Using maps of the area to determine a purchaser’s residence, the 95% confidence interval for the estimated fraction,  $p$ , of out of city sales was  $0.367 \pm 0.03$ . As the company had paid the city \$76,975, the corresponding confidence interval for the amount of overpayment is  $\$28,250 \pm 2310$ . Although accounting experts testified that sampling methods were an accepted method for conducting audits, the court did not accept the results because the sales tax return called for a computation for each sale. Subsequently, a complete audit was made, which yielded an overpayment of \$26,750. The complete audit cost \$3500. The sample cost slightly more than 10%

of the complete audit, substantiating its economic advantage as the cost-differential exceeds the difference between the sample estimate and the complete audit.

Samples were accepted in cases concerning whether the quality of goods met a standard specified in a law (Larsen, 1964). For example, the Food, Drug and Cosmetic Act and related laws allow the government to seize food or other goods that may be harmful to the public. Courts realized that one could not test every item in a large shipment, suspected of containing a high fraction of defective products, as the testing process was destructive. The case, *U.S. vs. 43 ½ Gross of Rubber Prophylactics* 65 F. Supp. 534 (D. Minn. 1946) is illustrative. The government condemned an entire shipment containing 22,464 condoms made by a firm after a 7.35% (30) of a sample of 408 tested had holes. As the devices were sold, at the time, for the purpose of preventing venereal disease, the court realized that testing rendered the condoms unfit for resale. Thus, the government was correct in relying on a sample. The court also found that condoms with a 7% chance of being defective would subject the consuming public to an increased risk of disease. For a fuller description of the case, see Gastwirth (1988).

### 3. The use of surveys and statistical studies in cases involving the Lanham Act

The Lanham Act, Sec. 43(a) 15 U.S.C. Sec. 1125(a), prohibits deceptive or misleading advertising and infringing on a trademark or patent. The infringer is liable for the estimated profit lost to the holder of the original trademark. These can range from token (\$1) to substantial (\$17 million) amounts (Carter and Remec, 1996). Diamond (2000) notes the plaintiff, typically the firm who believes its trademark is being infringed or its product unfairly denigrated in a competitor's advertising campaign, often asks courts for a preliminary injunction. Such a ruling bars the defendant from using the advertisement or selling the product in question while the judicial process is operating. The plaintiff must show that they are very likely to prevail at trial to obtain such an injunction. Survey evidence is used to demonstrate that potential consumers are likely to be misled by an advertisement or confused as to the actual maker of a product. In order to be persuasive in litigation a survey needs to be carefully designed to focus on the specific issues involved in order to isolate the effect of the defendant's alleged infringement.

The survey results submitted into evidence are rarely random samples of potential consumers but are market research surveys, typically taken in a number of malls. Therefore, it is important that the survey incorporate questions that identify the appropriate respondents, namely individuals who are potential consumers or purchasers of the product. As Diamond (2000) emphasizes the purpose of surveys in these cases is not simply to ascertain the beliefs of consumers but how the trademark or commercial influenced those beliefs. Thus, in deceptive advertising cases surveys should focus on learning how the content of a commercial changes or affects the respondents' perceptions of the product. Similarly, surveys submitted in trademark confusion cases should ascertain how respondents rely on the mark to identify the source of the product. To accomplish this, it is useful to incorporate a control group or some control questions that enable us

to estimate the effect of the advertisement or allegedly infringing (similar) trademark. For example, to ascertain whether an advertisement is misleading, one can randomly split the sample into two groups. One is shown the actual commercial while the other sees an edited version of the commercial one without the possibly misleading portion. Both groups are asked the same questions designed to determine whether they believe that the advertisement claimed that one product was superior. If the proportion of viewers of the advertisement with the questioned statements who believe that it claimed the new product was better is significantly higher than the corresponding proportion of the control group, then the difference between these proportions estimates the effect of the questionable part.

### 3.1. Cases concerning trademark infringement

The Lanham Act protects holders of registered trademarks from competitors using marks, labels or packaging that are likely to cause confusion amongst potential purchasers of the protected product. The law is designed to ensure that consumers are not confused about the connection of a product and its manufacturer. Thus, a mark that leads potential purchasers to associate another product with that produced by another manufacturer infringes on the original trademark. Infringement includes misleading consumers about the origin, sponsorship or approval of the goods at issue.

The fundamental question in trademark infringement actions is whether there exists a likelihood that an appreciable number of ordinarily prudent consumers will be misled or confused as to the source of the goods in question. While evidence of actual confusion of even a small number of individuals is helpful, it is not required. The test is *likelihood* of confusion. This criteria is used as holders of major trademarks will often attempt to enjoin the sale of infringing goods as soon as they find they are on the market. Courts assess a number of factors including the distinctiveness of the senior (original) mark, the similarity of the two marks, the similarity of the goods, the similarity of the locations where the two products are sold and the advertisements both parties create and the degree of confusion. The degree of care an ordinary consumer would be expected to take is considered, e.g., a prudent consumer would examine a high priced item with greater care than a souvenir T-shirt.

Carefully taken surveys of potential consumers are useful in ascertaining whether there is a likelihood of confusion. Diamond (2000) discusses the properties of a valid survey. We mention several of the important ones. The survey should attempt to obtain a representative sample of the target population, i.e., even if only mall shoppers can be sampled for practical reasons, then samples should be taken from malls in several locations at different times of the day. The respondents should be screened to ensure they are in the target population, typically potential purchasers. The screening questions should not convey the sponsor or purpose of the survey to avoid pre-conditioning the respondents. The questions should be clear and respondents should be allowed to say they do not know. The questions should not lead the respondents; in particular order effects can be minimized by having the questions go from the general to specific or by having different orders (rotation) given to different respondents. Finally, a control group or question is very useful to separate out the level of confusion due to the

alleged infringing mark (or trade dress) and the background rate due to pre-existing knowledge. McCarthy (1996) notes that the closer a survey mirrors the situation in which the ordinary person would encounter the trademark the greater the evidentiary value of the results.

While courts accept non-probability samples that satisfy the major criteria of reflecting the population of potential users of the plaintiff's product and interviewing them properly, they have rejected surveys that have major flaws. Thus, in *Exxon Corp. vs. Xoil Energy Resources* 552 F. Supp. 1008 (S.D.N.Y., 1981) a court did not accept a survey of 194 individuals interviewed at a single New York City train station as reflective of a national market. If a case concerns a smaller area such as a county where the two marks both exist, then the sample should be taken from that county. Courts do realize that it is difficult to obtain a survey of the entire target population and are more concerned with under-inclusion than slight over-inclusion. After all in the second situation, one can delete the "extra" observations. In *Brooks Shoe Mfg. vs. Suave Shoe Corp.* 533 F. Supp. 75 (S.D. Fla. 1981), a case concerning the public's recognition of a design trademark on a running shoe, the plaintiff only sampled spectators and participants at running events in the Washington–Baltimore area. The court held that this sample of individuals seriously involved in running would be more likely to recognize the design than the typical consumer of athletic shoes, leading to an overestimate of the strength of the design. A number of other flaws, including the inclusion of a *Brooks* shoe among those supposed to have been made by *Suave* and the survey itself are discussed in Gastwirth (1988, pp. 515–520). Diamond (2000) lists several cases where survey results were discounted because they did not incorporate a control group or control question.

An illustrative case involving a control group is *Indianapolis Colts, Inc. vs. Metropolitan Baltimore Football Club Inc.* 34 F. 3d 410 (7th Cir. 1994). After the Baltimore Colts moved to Indianapolis they kept their name. Nine years later, the Canadian Football League (CFL) granted Baltimore a license, that team also took the name, Colts. The plaintiff surveyed several hundred consumers in 24 malls around the country. After screening respondents to establish they were football fans and had purchased merchandise with team names on it, they were asked questions specific to the Baltimore CFL Colts, e.g., what sport they played, what league they were in, etc. Nearly 60% of the fans shown the Baltimore CFL Colts merchandise thought it referred to the U.S. team. A control group was asked about a team named the Baltimore Horses. A much smaller percentage of them thought they were an U.S. football team. As the city is the same, the difference, over 40%, reflects the percentage of consumers misled by the name of the team.

The opinion, by Judge Posner, noted that one could ask whether the name chosen for the control group was as attractive as the real one or whether the sampled universe should have been narrowed to football fans who had bought merchandise with football team names on it. Nonetheless, one cannot expect perfection and it seems unlikely that they caused a substantial over-estimate of confusion. This is a very important point that arises in many areas of litigation, e.g., anti-trust and equal employment. Gastwirth (1994) notes that one can always point to a potential flaw; however, one needs to ask whether it is sufficient to alter the ultimate conclusion.

A recent infringement case discussing the evaluation of potential flaws in a survey is *Sara Lee Corp. vs. Kayser–Roth Corp.* 38 USPQ2d 1449 (4th Cir. 1996). Sara Lee sells hosiery under the *L'eggs* trademark. Its most popular product is its Sheer Energy brand of pantyhose. Kayser–Roth, its major competitor, introduced a product line named *Leg Looks*. The opinion of Judge Hall notes that the goods in question are very similar and are sold and advertised in similar places. In addition to some evidence of actual confusion of consumers about the maker of *Leg Looks*, Sara Lee introduced several surveys indicating that about 30–40% of the consuming public was confused by the similarity of the two names. The district court had discounted the survey on the grounds that its reliability may have been questionable but details are not provided in the appellate opinion. From a scientific perspective, Judge Hall reviews a number of previous cases finding that courts have accepted survey confusion levels as low as 10–12% as sufficient, but lower levels, e.g., 7.6% insufficient, evidence of infringement. The opinion then notes that even if the true figure were just one-half of the survey estimate, it would indicate a legally significant degree of confusion.

Judge Hall's opinion is using a form of sensitivity analysis in this context. Essentially, the court found that whatever questions about the reliability were raised by Kayser–Roth, they were insufficient to reduce the level of confusion to a legally acceptable one. Sensitivity analysis, pioneered by Cornfield and developed by Rubin, Rosenbaum and Gastwirth (see Rosenbaum (1995) for early references and Gastwirth (1997) for its use in equal employment cases) examines the degree to which an omitted variable could explain an observed statistically significant association. In the context of estimation of a proportion in a survey, one can model the potential effect of a flaw and assess whether the imperfection could reduce that proportion to a legally negligible one. The issue is similar to assessing the effects of a possible explanatory variable that is omitted variable from the model (called “lurking variables” by Box et al. (1978) or measurement errors in the predictor variables, which arise often in econometric regression analysis (Theil, 1971; Dempster, 1988).

Trademark surveys do not have available a proper frame listing the universe of potential buyers. Rather, screening questions are utilized to determine whether someone is in that universe and mall surveys miss those shoppers who shop elsewhere. One can assess the potential impact of an imperfect frame on the estimated proportion of customers who are confused using the following method.

**Lemma 1.** *Let  $w$  denote the fraction of the actual sample who are in the appropriate universe and let  $\pi$  = the probability a true potential buyer would be confused and  $R\pi$ , the corresponding probability for other individuals, where  $R > 1$ . In order for a survey to obtain a fraction,  $s$ , of confused individuals when the true fraction is only  $\pi$ ,  $R$  and  $w$  must satisfy*

$$R \geq s/\pi \quad \text{and} \quad w \leq (1-s)/(1-\pi).$$

**Proof.** Note that  $s = w\pi + (1-w)R\pi$ . As  $R\pi$ , the probability that a individual who is not a potential buyer is confused, cannot exceed 1,  $s \leq w\pi + (1-w)$ . This implies the second inequality. Also,  $R - (R-1)w = s/\pi$ . As both  $w$  and  $R-1$  are positive, the first inequality holds.

*Comment:* The strength of the relationship, indicated by  $R$ , between the required level of extra confusion ( $R\pi$ ) of those sample members who are not truly potential purchasers and  $s/\pi$  is given by  $R = s/\pi + (R - 1)w$ . Thus, the required relative increase,  $R$ , in the probability of confusion of non-purchasers is greater than  $s/\pi$ . Solving for  $w$  yields  $w = (R - s/\pi)/(R - 1)$ . This implies that the closer  $R$  is to the minimum value,  $s/\pi$ , needed to explain the higher level of confusion found in the survey due to the inclusion of respondents who are not potential purchasers the smaller the fraction,  $w$ , of eligible respondents must be. This result enables us to assess whether a sample is sufficiently unrepresentative that its estimated level of confusion should be discounted.

To illustrate the use of Lemma 1 and the stronger result in the Comment, consider the 30–40% figure in *Sara Lee*. Suppose the true level of confusion amongst appropriate consumers was only 7.5% ( $\pi = 0.075$ ), a figure deemed too low to indicate confusion. We will use the lower 30% figure,  $s = 0.30$ , for the estimated proportion of confused customers in the survey as sampling error was not given for the 30–40% confused. The lemma tells us that  $R$ , the factor indicating how much more likely a non-potential purchaser will be confused as to the source of the product than a potential customer, is at least  $s/\pi = 0.30/0.075 = 4$ , or the confusion rate amongst non-customers is at least 30% and  $w \leq (1 - s)/(1 - \pi) = 0.7/0.925 = 0.757$ . Thus, at least one-quarter ( $(1 - w) = 1 - 0.757 = 0.243$ ), of those surveyed would have had to be non-customers in order for imperfections in the frame to create the observed level of confusion in the *Sara Lee* survey. More importantly, if we use the result in the Comment and assume that non-customers have a confusion level of 50% ( $R\pi = 0.50$ ), i.e.,  $R = 0.50/0.075 = 6.667$  or non-customers are over six times more likely to be confused than true potential purchasers, then  $w$  needs to be  $(R - s/\pi)/(R - 1) = 2.667/5.667 = 0.471$ . This implies that in order for a survey to have been so flawed that the true level of confusion was only 7.5%, over 50% of the respondents would have had to be inappropriate ( $1 - w = 1 - 0.471 > 0.50$ ) and they had to have a confusion rate of 50% ( $R\pi = 0.50$ ). By assuming realistic rates of confusion for non-customers and some plausible values for the fraction  $(1 - w)$  of the surveyed respondents who were inappropriately included, one can assess whether a somewhat imperfect sample could reduce the observed fraction of confused potential purchasers to a legally insignificant one. In the *Sara Lee* case, we realize that not only would a substantial fraction (50%) of the surveyed individuals need to be non-customers, they also would need to have a high (50%) confusion rate. Thus, Judge Hall's conclusions are strengthened by a formal sensitivity analysis.

Judge Hamilton's recent opinion in *Simon Property Group L.P. vs. mySimon, Inc.*, 104 F. Supp. 2d 1033 reviews many cases and carefully evaluates the plaintiff's proposed survey. In particular, flaws need to be severe to render the survey inadmissible as relatively minor flaws go to the weight of the evidence. The discussion contrasting the survey accepted in *Sands Taylor and Wood Co vs. Quaker Oats*, 978 F. 2d 947 (7th Cir. 1992) case and the one rejected in the current case is illuminating. In *Sands* the appeals court reversed a lower court's rejection of survey evidence that 24% of consumers shown a label of a product called "Thirst-Aid" thought it was produced by Gatorade. They were not shown a Gatorade mark or label so their responses were not influenced by the plaintiff's mark. In contrast, in *mySimon* respondents were shown both web-sites at a mall location.

Table 1

Data from the blind test comparison of Triumph and Merit brands

Results	Triumph much better than Merit	Triumph somewhat better	Triumph and Merit about the same	Triumph somewhat worse	Triumph much worse than Merit
Number	45	73	77	93	36
Percent	14	22	24	29	11

*Castrol vs. Pennzoil Co.* (987 F.2d 939 (3rd Cir. 1993) the plaintiff sued Pennzoil because it advertised that its oil outperformed any other brand against viscosity breakdown and thus offered better protection against engine breakdown. The opinion noted that the oils are intended to provide a protective film to minimize metal to metal contact and that they are designed to resist flow to maintain that protection. This resistance to flow is measured by its viscosity. The industry has established two tests for evaluating the viscosity of motor oils, the shear stability test and the high-temperature/high shear test. Both products passed the first test and the plaintiff's product actually was slightly superior on the second test to some of the defendant's. The defendant, however, claimed its products performed better on a third test based on measuring the percent viscosity loss. The trial court had found, however, that this third test did not measure viscosity breakdown. Indeed, the appellate opinion observed that this third test was never intended to compare the viscosity breakdown of oils of different types; rather it was designed as a quality control test for manufacturers to ensure that the quality of successive batches of oil were similar. This case highlights the importance of measuring the scientifically relevant variables. A related case, *Castrol vs. Quaker State Corp.* (977 F.2d (2nd Cir. 1992), found that Quaker State's advertisements claiming that because its oil had faster oiling time, i.e., reached distant engine parts faster than its competitor's products, it reduced engine wear were false. Indeed, in one of Quaker State's studies of 64 engine start-ups there was no significant difference in metal wear, although its oil was faster. The district court had issued an injunction stopping the advertisements, which was upheld by the appellate court.

We next turn to the second class of cases where the issue is whether an advertisement, while literally true, misleads a substantial fraction of consumers. The recent decision in *Clorox Co. Puerto Rico vs. Procter and Gamble Co.* (56 USPQ2d 1385, 1 Cir. 2000) notes that an advertisement's tendency to mislead a substantial portion of its intended audience is most often proved by consumer survey data. Indeed, the appellate courts have emphasized that the trial court's reaction is not especially relevant to the basic issue, which is what the people to whom the advertisement is directed believe it says. The opinion also noted that survey evidence is not required if the defendant "intentionally set out to deceive the public". For further description of the legal as well as statistical elements of a false advertising claim, see McCarthy (1996).

The decision in *Clorox* reversed a lower court's dismissal of its claim that the advertisements used by P&G in Puerto Rico to sell its Ace with whitener were misleading. After P&G's consumer studies indicated that an obstacle it faced with its product was the public's perception that chlorine bleach was necessary to get clothes white, it ran

commercials claiming “Compare with your detergent ... whiter is not possible”. *Clorox* had submitted a survey showing that 47% totally agreed and 20% somewhat agreed with the statement that ACE left clothes as white or whiter than if one uses bleach. The trial court decided that surveys should not be used “to determine the meaning of words or set the standard to which objectively verifiable claims must be held”. The appellate opinion noted that when a producer invites consumers to compare its product to others, it is making a specific and measurable claim. At trial the issues of whether the ad campaign was literally false or misleading will be given a full hearing. In addition to surveys, other scientific studies measuring whiteness may be submitted.

### 3.3. Trademark dilution cases

In 1995 Congress expanded trademark protection to include protection against similar marks on dissimilar products that would dilute the positive association consumers have developed with the original mark (Garcia, 1995; Bible, 1999). The role that surveys may have in this type of case remains uncertain as the Circuits have split as to whether evidence of actual dilution or a likelihood of dilution is required (Morrison, 2000). Surveys have greater relevance under the second criteria.

A recent case, *Eli Lilly Co. vs. Natural Answers, Inc.* 200 F.3d (7th Cir. 2000) affirmed a preliminary injunction against the use of “herbprozac” on a “mood elevating” herb supplement because the name would dilute the “Prozac” trademark registered by Lilly. The court noted that although the record did not contain evidence of actual confusion, the “cause dilution” requirement of the law is satisfied by proof of a likelihood of dilution. The opinion noted that requiring actual confusion would mean that senior mark holders would be prevented from filing suit before they were economically harmed. Moreover, a new company would be unable to request a declaratory judgment that their mark was sufficiently different from an existing one prior to expending the funds necessary to launch a new product.

In a classic article, Schechter (1927) observed that in modern economies, where products often pass through a chain of intermediaries before being purchased by the consumer, the value of a trademark is due to its uniqueness as identifying the quality of the product and its corresponding selling power. Thus, a mark can be diluted if the reputation of its holder may be affected by another manufacturer’s use of it. Although several states passed anti-dilution laws protecting marks, the Federal law only went into effect in 1996.

The new law (Bible, 1999) protects owners of famous marks and lists eight factors courts may consider. In particular, survey evidence can bear on assessing the degree of distinctiveness of the mark to the public and the degree of recognition of the mark in the trading areas and channels of trade used by the mark’s owner and the possibly infringing product. In dilution suits the issue is not whether consumers are confused as to the source of the product; rather it is whether the junior user’s mark has resulted in a “lessening of the capacity of a famous mark to identify and distinguish goods and services”. Thus, individuals exposed to both marks need to be surveyed to determine whether potential customers associate the infringing product as being made or distributed by the holder of the senior mark. As in surveys used in “likelihood of

confusion” cases, a comparison of the results of a survey of individuals unexposed to the junior party’s mark with the results of the survey of the exposed group enables one to show that the infringing mark has diluted the original one. This is important because one is studying whether there has been a loss of the original mark’s potential selling power or identity.

Survey evidence is not required in a dilution case as often the facts, especially the fame of the original mark and the intent of the manufacturer of the junior mark, of a case are quite clear. A survey was used successfully in *Wawa vs. Haaf* 40 U.S.P.Q. 2d (E.D. Pa.1996). The “Wawa” chain of convenience stores is well known in the northeast. The defendant opened a similar store in Pennsylvania using the name “HaHa”. A survey of the population in the neighborhood of the new store showed that 29% of those interviewed associated the new store with Wawa. The court found that this was sufficient to strengthen Wawa’s claim of trademark dilution and enjoined the defendant from using the “HaHa” mark.

The new law also raises some interesting questions as to the economic meaning and measurement of dilution. Hartman (1997) introduced the concept of “brand equity” as the value a trademark provides. It includes the perceptions and associations, presumably positive, consumers have experienced with the product. Bible (1999) and Simonson (1993) discuss the use of responses on a scale, say from 1 to 10, of consumer perceptions of the quality of the major brand to measure its “brand equity”. Loss of the power of a mark to uniquely identify a product has been assessed by a “cue recall” test, which asks consumers to identify the product lines manufactured under a given trade name. The proportion of consumers who list the junior party’s product reflects the potential for dilution. This quantity would be measured more accurately if there were a comparative figure. Firms that periodically survey consumers to measure their brand recognition might well incorporate a “cue recall” question as this would establish baseline figures both for the percentage of consumers who identify the product with the brand and the small percentage who misidentify it. This second percentage can be regarded as a control to which the proportion of respondents who identify a junior party’s mark in a subsequent dilution survey can be compared.

#### 4. Discussion

Survey evidence is a useful tool in legal cases, especially in the trademark protection area. Economists and statisticians interested in this area, especially if they plan to testify in court, should be aware of the recent Supreme Court decisions designed to ensure the reliability and validity of scientific evidence (Berger, 2000; Rosenblum, 2000) and recent revisions of the Federal Rules of Evidence. Trial judges are now responsible for assessing the reliability of the proposed expert testimony and the qualifications of the expert. Both are examined in light of their relevance to the issues in the case at hand. It is important for potential experts to understand the purpose of the survey in the context of the litigation, i.e., which parameter is it supposed to measure and design an appropriate study. A summary report should not only present the results, e.g., percent “confused” along with a confidence interval, it should contain a description of the

response rate, the method of data collection and a copy of the questions and interviewer instructions (Diamond, 2000).

## Acknowledgements

It is a pleasure to thank Dr. Joseph Cecil, Dr. Binbing Yu and both referees for several valuable suggestions and references. I also want to thank Prof. D. Slottje for encouraging me to write on this topic. This research was supported in part by a grant from the National Science Foundation.

## References

- Aitken, C.G.G., 2000. Interpretation of evidence and sample size determination. In: Gastwirth, J.L. (Ed.), *Statistical Science in the Courtroom*. Springer, New York, NY, pp. 1–24.
- Berger, M.A., 2000. The Supreme Court's trilogy on the admissibility of expert testimony. In: *Reference Manual on Scientific Evidence*, 2nd Edition. Federal Judicial Center, Washington, DC, pp. 9–38.
- Bible, P.M., 1999. Defining and quantifying dilution under the federal trademark Dilution Act of 1995: using survey evidence to show actual dilution. *University of Colorado Law Review* 70, 295–340.
- Box, G.E.P., Hunter, W.G., Hunter, J.S., 1978. *Statistics for Experimenters*. Wiley, New York, NY.
- Carter, A.W., Remec, G.M., 1996. Monetary awards for trademark infringement under the Lanham Act. *Trademark Reporter* 86, 464–484.
- Craswell, R., 1997. "Compared to What?" The use of control ads in deceptive advertising litigation. *Antitrust Law Journal* 65, 757–791.
- Dempster, A.P., 1988. Employment discrimination and statistical science (with discussion). *Statistical Science* 3, 149–195.
- Diamond, S.S., 1994. Reference guide on survey research. In: *Reference Manual on Scientific Evidence*, 1st Edition. Federal Judicial Center, Washington, DC, pp. 221–270.
- Diamond, S.S., 2000. Reference guide on survey research. In: *Reference Manual on Scientific Evidence*, 2nd Edition. Federal Judicial Center, Washington, DC, pp. 229–276.
- Garcia, J.A., 1995. Trademark dilution: eliminating confusion. *Trademark Reporter* 85, 489–524.
- Gastwirth, J.L., 1988. *Statistical Reasoning in Law and Public Policy*. Academic Press, New York, NY.
- Gastwirth, J.L., 1992. Statistical reasoning in the legal setting. *American Statistician* 46, 55–69.
- Gastwirth, J.L., 1994. Review of Diamond (1994). *Jurimetrics Journal* 36, 181–191.
- Gastwirth, J.L., 1997. Statistical evidence in discrimination cases. *Journal of the Royal Statistical Society (A)* 160, 289–303.
- Hall, R.E., Lazear, V.A., 2000. Reference guide on estimation of economic losses and damages awards. In: *Reference Manual on Scientific Evidence*, 2nd Edition. Federal Judicial Center, Washington, DC, pp. 277–332.
- Handler, A.M., McCarthy, E.P., 1996. Conopco—Consumer confusion remains the touchstone of trade dress infringement. *Trademark Reporter* 86, 595–610.
- Hartman, S., 1997. Brand equity impairment—the meaning of dilution. *Trademark Reporter* 87, 418–436.
- Heiner, K.W., Wagner, N.L., Fried, A.C., 1984. Successfully using statistics to determine damages in fiscal audits. *Jurimetrics Journal* 24, 273–281.
- Izenman, A.J., 2000. Statistical issues in the application of the federal sentencing guidelines in drug, pornography and fraud cases. In: Gastwirth, J.L. (Ed.), *Statistical Science in the Courtroom*. Springer, New York, NY, pp. 25–49.
- Kaye, D.H., Freedman, D.A., 2000. Reference guide on statistics. In: *Reference Manual on Scientific Evidence*, 2nd Edition. Federal Judicial Center, Washington, DC, pp. 83–178.
- King, B.F., 1986. Statistics in anti-trust cases. In: Morris, H.D., Stephen, E.F., Joseph, B.K. (Eds.), *Statistics and the Law*. Wiley, New York, NY.

- Landes, W., Posner, R.A., 1988. The economics of trademark law. *Trademark Reporter* 78, 267–306.
- Larsen, K.H., 1964. Admissibility in evidence of sample or samples of articles or substance of which the quality, condition, or the like is involved in litigation. *American Law Reports* 2d 95, 681–712.
- McCarthy, J.J., 1996. Lanham Act 43(a): the giant is now wide awake. *Law and Contemporary Problems* 59, 45–74.
- Morrison, T.C., 2000. Federal dilution act at 5: a qualified success story. *National Law Journal* C6.
- Palladino, V.F., 1994. Surveying secondary meaning. *Trademark Reporter* 84, 155–185.
- Rosenbaum, P.R., 1995. *Observational Studies*. Springer, New York, NY.
- Rosenblum, M., 2000. On the evolution of analytic proof, statistics, and the use of experts in EEO litigation. In: Gastwirth, J.L. (Ed.), *Statistical Science in the Courtroom*. Springer, New York, NY, pp. 161–194.
- Schechter, F.I., 1927. The rational basis of trademark protection. *Harvard Law Review* 40, 813–833.
- Simonson, A.F., 1993. How and when do trademarks dilute: a behavioral framework to judge “likelihood” of dilution. *Trademark Reporter* 83, 149–174.
- Sprohls, R.C., 1956. The admissibility of sample data into a court of law: a case history. *UCLA Law Review* 4, 222–232.
- Theil, H., 1971. *Principles of Econometrics*. Wiley, New York, NY.
- Walker, L., Monahan, J., 1998. Sampling damages. *Iowa Law Review* 83, 545–568.