

THE ROLE OF CONSUMER SURVEYS IN TRADEMARK INFRINGEMENT: EMPIRICAL EVIDENCE FROM THE FEDERAL COURTS

Robert C. Bird* & Joel H. Steckel**

With millions, perhaps billions, of dollars at stake in the value of a brand,¹ brand equity can be one of the most important assets in a firm's portfolio.² Unfortunately, brand equity is an asset that is uniquely vulnerable to harm. Firms can lose the strength, and thus the selling power, of those brands through the ordinary course of business.³ Even a firm's

* Associate Professor and Northeast Utilities Chair in Business Ethics, School of Business, University of Connecticut. We thank Meredith Long and Michael Thomason for valuable research assistance. We also thank Barton Beebe for sharing his data with us.

** Professor of Marketing, Stern School of Business, New York University.

1. A ranking by *Bloomberg Businessweek* of the top one hundred global brands reveals estimated values for many top brands in excess of ten billion dollars. BLOOMBERG BUSINESSWEEK, 100 BEST GLOBAL BRANDS 1, available at, http://www.businessweek.com/interactive_reports/best_global_brands_2009.html. The top hundred most valuable global brands are worth more than two trillion dollars. See KEN SCHEPT, BRANDZ TOP 100 MOST VALUABLE GLOBAL BRANDS (MilwardBrown 2010), available at http://c1547732.cdn.cloudfiles.rackspacecloud.com/BrandZ_Top100_2010.pdf.

2. See, e.g., Robert C. Bird, *Pathways of Legal Strategy*, 14 STAN. J.L. BUS. & FIN. 1, 28 (2008) (highlighting Google's use of brand equity for competitive advantage); Sonia K. Katyal, *Stealth Marketing and Antibranding: The Love that Dare Not Speak Its Name*, 58 BUFF. L. REV. 795, 804 (2010) (noting that the value of a firm's brand equity can exceed the book value ascribed to its product). See generally Jerre B. Swann, *An Interdisciplinary Approach to Brand Strength*, 96 TRADEMARK REP. 943, 945-54 (2006) (describing the value of brand equity); Jerre B. Swann & Theodore H. Davis, Jr., *Dilution, An Idea Whose Time Has Gone: Brand Equity as Protectable Property, The New/Old Paradigm*, 1 J. INTELL. PROP. L. 219, 219-20 (1994) (positing that brand equity is a property right that should be entitled to its own protection); Judith Zaichkowsky, *Strategies for Distinctive Brands*, 17 J. BRAND MGMT. 548, 548 (2010) (stating the generally accepted principle that brand equity is essential for successful marketing). For in-depth work on brands from a leading author on the issue, see DAVID A. AAKER, BRAND RELEVANCE: MAKING COMPETITORS IRRELEVANT (2010); DAVID A. AAKER, BUILDING STRONG BRANDS (1995); DAVID A. AAKER, MANAGING BRAND EQUITY: CAPITALIZING ON THE VALUE OF A BRAND NAME (1991).

3. Negative brand effects can occur in a variety of contexts. See, e.g., Luis J. Diaz & Patrick C. Dunican Jr., *Ending the Revolving Door Syndrome in Law*, 41 SETON HALL L.

own misguided behavior through mismanagement or over-popularization of the brand can dilute the influence of even the most elite of brand names.⁴ Damage to a brand can be so devastating that a firm may wish to deliberately and permanently disassociate itself from its meaning to consumers.⁵

Brand equity is no less vulnerable to the actions of competitors. For example, if one firm uses another's famous brand in a way that would confuse consumers, the value of the brand to its established owner would suffer.⁶ In such circumstances the owner of the brand could sue the challenger for trademark infringement in violation of the Lanham Act, which prohibits an unauthorized user of a trademark from using it in a way that "is likely to cause confusion, or to cause mistake, or to deceive"⁷

Litigants protecting their marks have traditionally presented three different types of evidence to prove in court a likelihood of confusion by a

REV. 947, 962 (2011) (discussing brand reputation for professional service firms); Steve Hartman, *Brand Equity Impairment—The Meaning of Dilution*, 87 TRADEMARK REP. 418, 428–33 (1997) (describing brand equity impact in a variety of consumer contexts); Anupam Jaju et al., *Consumer Evaluations of Corporate Brand Redeployments*, 34 J. ACAD. MARKETING SCI. 206, 206 (2006) (discussing merger and acquisition activities).

4. Gucci diluted its elite brand by attaching its name to 22,000 widely different products, some of which were of low quality and of questionable connection to the Gucci image. See Kevin Lane Keller, *Managing the Growth Tradeoff: Challenges and Opportunities in Luxury Branding*, 16 J. BRAND MGMT. 290, 298 (2008). See generally Barbara Loken & Deborah Roedder John, *Diluting Brand Beliefs: When Do Brand Extensions Have a Negative Impact?*, 57 J. MARKETING 71 (1993) (examining situations in which brand extensions are likely to dilute beliefs associated with family brand name); Pavel Štrach & André M. Everett, *Brand Corrosion: Mass-Marketing's Threat to Luxury Automobile Brands After Merger and Acquisition*, 15 J. PRODUCT & BRAND MGMT. 106, 115–16 (2006) (finding dilution in luxury brands subsequent to overextension of models and production quantities).

5. See Aaron Perzanowski, *Unbranding, Confusion, and Deception*, 24 HARV. J.L. & TECH. 1, 2 (2010) ("When a brand suffers from strong negative consumer perceptions, it transforms from a valuable asset to a major liability. Faced with the reality of an irreparably damaged brand, many firms understandably seek a fresh start.").

6. In trademark disputes, the firm that has established rights to the trademark and asserts that a rival is engaging in potentially infringing behavior is known as the senior user of the trademark. The mark itself is frequently called a senior mark. Correspondingly, the company that is being challenged for an allegedly infringing use is known as the junior user, and its challenged mark is known as the junior mark. See 1 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 2:10 (4th ed. 2012) ("[W]hen a junior user uses a mark similar to a senior user's mark such that there is a likelihood of confusion, this is infringement which may be enjoined. Thus, the scope of exclusivity of a trademark is coextensive with the prevention of confusion of customers."); Thomas R. Lee, *Demystifying Dilution*, 84 B.U. L. REV. 859, 885–86 (2004) ("Dilution occurs . . . whenever a junior mark interferes with a trademark's ability to function as an identifier of a unique source of goods.").

7. 15 U.S.C § 1114 (2005).

rival.⁸ Expert witnesses can testify about their beliefs regarding confusion,⁹ but such testimony may not represent the consumer's state of mind and can devolve into a "battle of the experts" between hired guns paid to support a particular position.¹⁰ Litigants also make visual comparisons between marks,¹¹ but such comparisons represent only a fraction of what consumers might discern in order to determine whether a trademark comes from a particular source.

The third type of evidence, and the one that has intrigued courts and scholars for decades, is the consumer survey.¹² A consumer survey is an instrument used to gather data on the beliefs and attitudes of consumers

8. 4 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at § 23:2.50.

9. 4 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at § 23:2.75; Michael J. Allen, *The Role of Actual Confusion Evidence in Federal Trademark Infringement Litigation*, 16 CAMPBELL L. REV. 19, 28 (1994) ("The best evidence of actual confusion is evidence of individuals who have purchased, as a result of similar trademarks, the product of one of the parties under the mistaken belief that they are purchasing the other party's product."); Keith M. Stolte, *Remedying Judicial Limitations on Trademark Remedies: Monetary Relief Should Not Require Proof of Actual Confusion*, 75 DENV. U. L. REV. 229, 235–36 (1997) (stating that consumer testimony relating to purchases made as a result of confusion is strong evidence of actual confusion).

10. Alison M. Andrews, Note, *Implied Misrepresentations in Advertisements Under Section 43(a) of the Lanham Trademark Act: American Home Products Corp. v. Johnson & Johnson*, 47 ALB. L. REV. 97, 130 (1982) (citing the landmark survey case of *Zippo Mfg. Co. v. Rogers Imps., Inc.*, 216 F. Supp. 670, 684 (S.D.N.Y. 1963)). Judge Richard Posner has also remarked on the problem of expert testimony in the context of trademark litigation:

Many experts are willing for a generous (and sometimes for a modest) fee to bend their science in the direction from which their fee is coming. The constraints that the market in consultant services for lawyers places on this sort of behavior are weak, as shown by the fact that both experts in this case were hired and, we have no doubt, generously remunerated even though both have been criticized in previous judicial opinions. The judicial constraints on tendentious expert testimony are inherently weak because judges (and even more so juries, though that is not an issue in a trademark case) lack training or experience in the relevant fields of expert knowledge. But that is the system we have.

Indianapolis Colts, Inc. v. Metro. Balt. Football Club Ltd. P'ship, 34 F.3d 410, 415 (7th Cir. 1994).

11. 4 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at § 23:2.50.

12. See, e.g., Robert C. Sorensen & Theodore C. Sorensen, *Responding to Objections Against the Use of Opinion Survey Findings in the Courts*, 20 J. MARKETING 133 (1955) (arguing in 1955 that current methods of opinion research were sufficiently developed to make a definite contribution to judicial process); Note, *Public Opinion Surveys as Evidence: The Pollsters Go to Court*, 66 HARV. L. REV. 498 (1953) (examining evidential and practical problems involved in using surveys in litigation).

towards products, concepts, or names.¹³ This data is gathered through a variety of means, including consumer queries through shopping malls, telephone contacts, and the Internet.¹⁴ Litigants use such surveys to convince a court that consumer confusion exists (or does not exist) between trademarks in lawsuits alleging trademark infringement.¹⁵ Consumer surveys provide direct evidence about consumer perceptions that expert testimony lacks and elicit multifaceted information about perceptions that mere visual comparison does not provide.

It is no wonder, then, that both courts and commentators hold consumer surveys in high esteem. Surveys are considered to have both widespread acceptance and vital influence in trademark infringement cases.¹⁶ Courts have on occasion even faulted litigants for not bringing a survey¹⁷ and expressed little sympathy even when a small company cannot bear the high costs of administering one.¹⁸ While courts instruct that surveys are not essential, some do not hesitate to remind litigants that consumer surveys are the most direct method of showing a likelihood of

13. E.g., Robert C. Bird, *Streamlining Consumer Survey Analysis: An Examination of the Concept of Universe in Consumer Surveys Offered in Intellectual Property Litigation*, 88 TRADEMARK REP. 269, 270 (1998) (“A consumer survey is a scientific method of presenting evidence of mental associations of a given group of people by asking a representative sample of the relevant target group. In other words, surveys gather data on the attitudes and beliefs of consumers towards a given product, name, or concept.”) (footnote omitted); Henry D. Ostberg, *Response to the Article Entitled, “A ‘Reading’ Test or a ‘Memory’ Test: Which Survey Methodology is Correct?”*, 95 TRADEMARK REP. 1446, 1446 (2005) (“Consumer surveys used as evidence in trademark litigation frequently involve showing respondents one or more products, advertisements or marks and asking the respondents for their perceptions and beliefs about these stimuli.”).

14. Bird, *supra* note 13, at 270.

15. See, e.g., K.J. Greene, *Abusive Trademark Litigation and the Incredible Shrinking Confusion Doctrine—Trademark Abuse in the Context of Entertainment Media and Cyberspace*, 27 HARV. J.L. & PUB. POL’Y 609, 620 (2004) (arguing that such reliance makes trademark cases fact-driven, and thus judicial determinations of likelihood of confusion are necessarily subjective and impressionistic).

16. See, e.g., Gabriel M. Gelb & Betsy D. Gelb, *Internet Surveys for Trademark Litigation: Ready or Not, Here They Come*, 97 TRADEMARK REP. 1073, 1075 (2007) (noting a general increase of survey use in litigation since the Supreme Court’s decision in *Daubert*); Michael Rappeport, *Litigation Surveys—Social “Science” as Evidence*, 92 TRADEMARK REP. 957, 957 (2002) (stating that survey evidence is necessary in some litigation under the Lanham Act); Itamar Simonson, *The Effect of Survey Method on Likelihood of Confusion Estimates: Conceptual Analysis and Empirical Test*, 83 TRADEMARK REP. 364, 364 (1993) (supporting the routine implementation and resulting importance of survey use).

17. See *Gimix, Inc. v. JS&A Group, Inc.*, No. 80 C 6592, 1982 WL 52164, at *1006 (N.D. Ill. Jan. 13, 1982), *aff’d*, 699 F.2d 901 (7th Cir. 1983) (admonishing both sides for failing to produce consumer surveys).

18. *Sports Traveler, Inc. v. Advance Magazine Publishers, Inc.*, 25 F. Supp. 2d 154, 164 (S.D.N.Y. 1998).

confusion among consumers in trademark infringement cases.¹⁹ Consumer surveys seem to matter.

However, in spite of the hundreds of articles and court cases that analyze consumer surveys, very little data-driven evidence exists to show that consumer surveys are actually widespread, influential, and important. The literature suffers from a lack of an empirical understanding about how federal courts actually treat consumer surveys, in spite of what scholars and judges write about surveys in publications and judicial opinions. Addressing this issue is not just yet another attempt to fill a scholarly gap, but also has great practical implications for firms faced with defending their multi-million or billion-dollar brands in court. Consumer surveys are expensive to create and time consuming to administer.²⁰ Thus, developing survey evidence can drain a firm's limited resources.

We present an empirical study investigating the role of consumer surveys in federal courts by examining more than five hundred court opinions over a seven-year period. Using these data, this article has three objectives. First, we assess how often consumer surveys are actually used in federal trademark infringement cases, notwithstanding the anecdotal assessment of survey use as pervasive. Second, we contemplate the rarely considered question of whether submission by a plaintiff or defendant influences the treatment of the survey by the court. In theory, it should make no difference at all, but the data have the real potential to prove otherwise. Third, we examine whether the strength of the underlying evidence in trademark cases impacts the power of the survey. Consumer surveys could be equally helpful or most influential when other evidence is absent. From these three objectives, we conclude that survey evidence is

19. *E.g.*, *Tone Bros. v. Sysco Corp.*, 28 F.3d 1192, 1204 (Fed. Cir. 1994) (quoting *Co-Rect Prods. v. Marvy! Adver. Photography, Inc.*, 780 F.2d 1324, 1333 n.9 (8th Cir. 1985)) (“Consumer surveys are recognized by several circuits as the most direct and persuasive evidence of secondary meaning.”); *see also* *Checkpoint Sys., Inc. v. Check Point Software Techs., Inc.*, 269 F.3d 270, 282 n.10 (3d Cir. 2001) (noting that surveys are not essential where other evidence of confusion exists).

20. *See* *Nabisco, Inc. v. PF Brands, Inc.*, 191 F.3d 208, 224 (2d Cir. 1999) (noting the monetary and time investment in consumer surveys); Dan Sarel & Howard Marmorstein, *The Effect of Consumer Surveys and Actual Confusion Evidence in Trademark Litigation: An Empirical Assessment*, 99 TRADEMARK REP. 1416, 1416 (2009) (“Obtaining any data from consumers, particularly when a properly conducted survey is entailed, is a complex, time-consuming, and expensive process.”); *see also* *Lon Tai Shing Co., Ltd. v. Kotch + Lowy*, No. 90 Civ. 4464(DNE), 1991 WL 170734, at *19 n.15 (S.D.N.Y. June 20, 1991) (claiming that in 1990 surveys could cost as much as \$65,000); Robert C. Bird, *The Impact of the Moseley Decision on Trademark Dilution Law*, 26 J. PUB. POL’Y & MARKETING 102, 104 (2007) (stating that surveys can cost between \$50,000 and \$100,000); Rebecca Tushnet, *Running the Gamut From A to B: Federal Trademark and False Advertising Law*, 159 U. PA. L. REV. 1305, 1339 (2011) (noting that consumer surveys are prohibitively expensive for smaller businesses).

used infrequently, treated subjectively, and has the potential to be either dispositive or useless depending on the context of the underlying evidence. We complete our article by explaining how firms can use our results to calculate mathematically whether developing a consumer survey is worthwhile given the firm's status as plaintiff or defendant and the strength of the underlying evidence.

Part I of this article reviews the Lanham Act, explores the survey literature, and highlights the reception given to consumer surveys by federal courts. Part II introduces our dataset and its method of analysis. Part III presents the results of our empirical testing and explains our cost-benefit analysis for choosing whether to create survey evidence. Part IV briefly presents policy implications and limitations of our study. Part V concludes.

I. TRADEMARK INFRINGEMENT AND CONSUMER SURVEYS

A. *The Lanham Act*

A trademark is a word or design used on an article of merchandise to identify it as the product of a particular manufacturer and to distinguish it from others.²¹ Trademarks are said to facilitate the transmission of accurate information and protect the consumer from confusion as to the source of a given product.²² Trademarks also establish a product's distinctiveness from its competitors, signal quality or other positive attributes, and serve as promotional tools.²³ Trademarks prevent consumers from being less able to

21. 15 U.S.C. § 1127 (2006) ("The term 'trademark' includes any word, name, symbol, or device . . . used by a person . . . to identify and distinguish his or her goods . . ."); 1 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at § 3:1 (explaining that trademarks "identify the source of one seller's goods and distinguish that source from other sources."); Vincent-Wayne Mitchell & Íde Kearney, *A Critique of Legal Measures of Brand Confusion*, 11 J. PRODUCT & BRAND MGMT. 357, 359 (2002).

22. See Robert G. Bone, *Hunting Goodwill: A History of the Concept of Goodwill in Trademark Law*, 86 B.U. L. REV. 547, 556 (2006) (noting that trademark law is designed to enable reliable transmission of information); Misha Gregory Macaw, *Google, Inc. v. American Blind & Wallpaper Factory, Inc.: A Justification for the Use of Trademarks as Keywords to Trigger Paid Advertising Placements in Internet Search Engine Results*, 32 RUTGERS COMPUTER & TECH. L.J. 1, 15 (2005) ("[T]rademarks . . . accurately reflect the producer associated with any individual product and consumers will be able to safely rely on trademarks as information sources in their purchasing decisions.").

23. Jacob Jacoby & Maureen Morrin, "Not Manufactured or Authorized by . . .": *Recent Federal Cases Involving Trademark Disclaimers*, 17 J. PUB. POL'Y & MARKETING 97, 97 (1998); Charles R. Taylor & Michael G. Walsh, *Legal Strategies for Protecting Brands from Genericide: Recent Trends in Evidence Weighted in Court Cases*, 21 J. PUB. POL'Y & MARKETING 160, 160 (2002).

distinguish desired products from their competitors.²⁴

Trademarks contribute to economic efficiency by reducing consumer search costs.²⁵ A consumer can look to trademarks as shorthand indicators of quality, prestige, or product attributes.²⁶ Producers benefit because they can invest in building goodwill with the confidence that others will not appropriate it.²⁷ Consumers benefit because trademarks relieve the burden of having to do exhaustive research about a product's features.²⁸ While marketers do not speak of trademarks *per se*, they attribute these same

24. Others characterize the prevention of consumer confusion as the primary goal of trademark infringement protection. *E.g.*, Harriette K. Dorsen, *Satiric Appropriation and the Law of Libel, Trademark, and Copyright: Remedies Without Wrongs*, 65 B.U. L. REV. 923, 940 (1985) ("The goal of the legislation is not to protect the business enterprise's trademark from encroachment, but to protect consumers from confusion . . ."); F. T. Alexandra Mahaney, Comment, *Incontestability: The Park 'N Fly Decision*, 33 UCLA L. REV. 1149, 1151–52 (1986) ("A trademark generally performs four functions deserving of protection: (1) it identifies one seller's product and distinguishes a product from similar products sold by others; (2) it signifies that all goods bearing the same trademark come from a single source; (3) it signifies that all goods bearing the same trademark are of equal quality; and (4) it is a prime instrument in the advertisement and sale of goods") (footnote omitted); Kenneth L. Port, *The Congressional Expansion of American Trademark Law: A Civil Law System in the Making*, 35 WAKE FOREST L. REV. 827, 896 (2000) ("[T]he principal legal goal of the system is to protect the consumer from confusion.") (footnote omitted); *see also* Inwood Labs., Inc. v. Ives Labs., Inc., 456 U.S. 844, 854 n.14 (1982) (noting that the two goals of the Lanham Act are to protect trademark owners and to protect consumers from confusing goods made by competing manufacture).

25. *See* Bone, *supra* note 22, at 555 (noting that trademarks reduce consumer search costs); *see also* Ty Inc. v. Perryman, 306 F.3d 509, 510 (7th Cir. 2002) ("The fundamental purpose of a trademark is to reduce consumer search costs . . ."); Macaw, *supra* note 22, at 15 (claiming that search costs are reduced when there is no source confusion).

26. *See* Robert G. Bone, *Enforcement Costs and Trademark Puzzles*, 90 VA. L. REV. 2099, 2173 (2004) (claiming that trademarks protect a product's "goodwill" or prestige value); Bone, *supra* note 22, at 554 (describing the use of trademarks as indicators); Daniel J. Gervais, *Intellectual Property, Trade & Development: The State of Play*, 74 FORDHAM L. REV. 505, 521 (2005) (stating that trademarks indicate the source of the good and can indicate the desired level of quality); Clarisa Long, *Dilution*, 106 COLUM. L. REV. 1029, 1034 (2006) (citing the classical view that a trademark is a proxy for certain product features).

27. *See* Gerard N. Magliocca, *One and Inseparable: Dilution and Infringement in Trademark Law*, 85 MINN. L. REV. 949, 958 (2001) (asserting that trademarks "encourage producers to invest" knowing that their investment cannot be "poach[ed]" by competitors); *cf.* Ann Bartow, *Likelihood of Confusion*, 41 SAN DIEGO L. REV. 721, 752 (2004) (arguing that unprotected marks would diminish the goodwill associated with the mark).

28. *See* Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 163–64 (1995) (reasoning that trademark law "reduce[s] the customer's costs of shopping and making purchasing decisions . . ." (quoting 1 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 2.01(2), at 2–3 (3d ed. 1994))); Stacey L. Dogan & Mark A. Lemley, *Trademarks and Consumer Search Costs on the Internet*, 41 HOUS. L. REV. 777, 778 (2004) (stating that trademark law can improve the flow of information in markets).

functions to brands.²⁹ Indeed, trademarks are essentially what marketers refer to as branding elements, the most salient being the brand name.³⁰

The statute most responsible for protecting trademarks against confusion is the Lanham Act of 1946.³¹ The Act prohibits the unauthorized use of a registered mark in a fashion that is “likely to cause confusion, or to cause mistake, or to deceive.”³² The goals of the Lanham Act are to provide a cause of action against those who use deceptive or misleading marks,³³ protect mark holders against unfair competition,³⁴ protect the ability of consumers to distinguish among competing producers, and avoid confusion about the potential source of a product.³⁵ Trademark owners

29. See Joel H. Steckel et al., *Dilution Through the Looking Glass: A Marketing Look at the Trademark Dilution Revision Act of 2005*, 96 TRADEMARK REP. 616, 621 (2006) (“[I]t goes without saying that much of what marketers have written about brands applies to trademarks.”). This does not mean that legal trademark policy and marketing practice always coordinate harmoniously. One tension is that marketers often want to make their brands easier to interpret and thus more descriptive. This goal, however, impedes the lawyer’s goal of protecting the trademark rights that identify those same brands. Trademarks that are more descriptive in nature have a greater burden to show secondary meaning than their less descriptive (and more fanciful) counterparts. See *id.* at 622 n.26.

30. E.g. Keller, *supra* note 4, at 292 (noting the importance of trademarks in protecting brand image).

31. Pub. L. No. 79-489, 60 Stat. 427 (1946) (codified as amended at 15 U.S.C. §§ 1051-1141n (2006)); Matthew D. Bunker, *You Can’t Handle the Truth (In Music): Does the Lanham Act Preempt State “Truth in Music” Laws?*, 16 COMM. L. & POL’Y 1, 2 (2011) (describing the Lanham Act as the “primary source of federal trademark protection.”); Edward T. Saadi, *Sound Recordings Need Sound Protection*, 5 TEX. INTELL. PROP. L.J. 333, 352 n.90 (1997) (“The Lanham Act is the primary federal trademark and unfair competition statute.”); Anne M. Mellen, Comment, *Awarding Attorneys’ Fees Under the Lanham Act: Egregious Litigation Conduct in the “Exceptional” Case*, 74 U. CIN. L. REV. 1111, 1113 (2006) (describing the primacy of the Lanham Act); see also Marlene B. Hanson & W. Casey Walls, *Protecting Trademark Good Will: The Case for a Federal Standard of Misappropriation*, 81 TRADEMARK REP. 480, 501 (1990) (discussing how state antidilution laws supplement the Lanham Act).

32. 15 U.S.C. § 1114(1)(a) (2006).

33. 15 U.S.C. § 1127 (2006) (“The intent of this Act is to . . . mak[e] actionable the deceptive and misleading use of marks in such commerce; to protect registered marks used in such commerce from interference by State, or territorial legislation; to protect persons engaged in such commerce against unfair competition”); see also Jennifer E. Rothman, *Initial Interest Confusion: Standing at the Crossroads of Trademark Law*, 27 CARDOZO L. REV. 105, 123 (2005) (noting that one purpose of Lanham Act is to regulate deceptive and misleading uses).

34. 15 U.S.C. § 1127 (2006).

35. *Park ’N Fly, Inc. v. Dollar Park & Fly, Inc.*, 469 U.S. 189, 198 (1985) (stating that the Lanham Act helps consumers distinguish among competing products); Peter W. Smith, Note, *Trademarks, Parody, and Consumer Confusion: A Workable Lanham Act Infringement Standard*, 12 CARDOZO L. REV. 1525, 1529 (1991) (“One of the principal goals of the Lanham Act is to protect consumers from confusion and deception in the marketplace.”).

who can successfully prove their case in court can obtain an injunction against the infringer to prevent the use from continuing.³⁶

Typically, the senior user of the mark is the plaintiff and alleges that the junior user's (i.e., the defendant's) use of the same or a similar mark creates a likelihood of confusion.³⁷ A small number of declaratory judgment cases reverse the typical role of plaintiff and defendant in that the plaintiff in those cases is the junior user asking the court to declare that a mark does not infringe.³⁸ In these declaratory judgment cases, the defendant (i.e., the senior user) often responds with a trademark infringement counterclaim and endeavors to show that a likelihood of confusion of its trademark has indeed occurred by the junior user.³⁹

A single statutory framework governs how a likelihood-of-confusion of relevant consumers is proven and remedied in federal court.⁴⁰ Yet, considerable variance exists as to how this framework is interpreted and applied. In 1961, the landmark case of *Polaroid Corp. v. Polarad Electronics Corp.* established a list of factors that a court should consider when determining whether a likelihood of confusion exists.⁴¹ The Second Circuit listed eight factors relevant for determining consumer confusion: (1) the strength of the plaintiff's mark; (2) the degree of similarity between plaintiff's and defendant's marks; (3) the proximity of the products; (4) the likelihood that plaintiff will bridge the gap; (5) actual confusion; (6) the

36. 15 U.S.C. § 1114(2) (2006).

37. See Kathleen B. McCabe, Note, Dilution-by-Blurring: A Theory Caught in the Shadow of Trademark Infringement, 68 *FORDHAM L. REV.* 1827, 1828 (2000) (asserting that the traditional trademark infringement claim occurs when a senior user asserts her trademark is misappropriated by the junior user in order to confuse consumers about the source of the good); 4 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at §§ 24:1 (describing the typical process for trademark infringement claims); Michael J. Meurer, *Inventors, Entrepreneurs, and Intellectual Property Law*, 45 *HOUS. L. REV.* 1201, 1204 n.13 (2008) ("The typical trademark infringement case involves a claim of direct confusion in which a defendant (the 'junior' or second user) exploits the goodwill of an established 'senior' user by selling products with a trademark so similar to the established company's trademark that consumers mistakenly believe the junior user's products come from the senior user.").

38. See 6 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at §§ 32:50, :57 (describing the reversal of roles that occurs in a small number of cases).

39. When we built our dataset (to be discussed shortly), we consistently treated the senior user as the plaintiff and the junior user as the defendant, even in declaratory judgment cases, to account for the typical incentives of the parties.

40. 15 U.S.C. § 1125(a)(1) (2006); 5 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, § 27:13; Sheldon W. Halpern, *A High Likelihood of Confusion: Wal-Mart, Traffix, Moseley, and Dastar—The Supreme Court's New Trademark Jurisprudence*, 61 *N.Y.U. ANN. SURV. AM. L.* 237, 248 (2005) (discussing the Lanham Act and the "peculiar federal scheme" for remedies).

41. 287 F.2d 492, 495 (2d Cir.), *cert. denied*, 368 U.S. 820 (1961).

defendant's good faith in adopting the mark; (7) the quality of defendant's product; and (8) the sophistication of the buyers.⁴²

The thirteen federal circuits, being largely independent operating jurisdictions, developed their own equivalents of *Polaroid* factors for proving likelihood of confusion.⁴³ Although some tests are similar to one another, considerable variance exists among circuits.⁴⁴ While some circuits list at least ten factors, other circuits only articulate six or seven.⁴⁵ While all circuits consider the strength of the plaintiff's mark, the similarity of the two marks, and the defendant's intent, only six circuits consider the similarity of sales facilities and only ten circuits consider the sophistication of the consumer in determining the likelihood of confusion.⁴⁶ Furthermore, four circuits do not consider the similarity of advertising and marketing methods between the litigants.⁴⁷ The result is that considerable variance exists among circuits about the proof necessary to show likelihood of confusion.⁴⁸ Table 1 lists the various *Polaroid* factors used by the different

42. *Id.* The court also acknowledged that this list was not exhaustive and that other variables may need to be considered. *Id.*

43. Likelihood-of-confusion tests vary significantly among circuits. *See Boston Duck Tours, LP v. Super Duck Tours, LLC*, 531 F.3d 1, 15 (1st Cir. 2008) (citing *Pignons S.A. de Mecanique de Precision v. Polaroid Corp.*, 657 F.2d 482, 487 (1st Cir. 1981)) (eight factors); *Frisch's Rests., Inc. v. Elby's Big Boy of Steubenville, Inc.*, 670 F.2d 642, 648 (6th Cir. 1982) (citing *AMF Inc. v. Sleekcraft Boats*, 599 F.2d 341, 348–49 (9th Cir. 1979)) (eight factors); *AMF*, 599 F.2d at 348–49 (eight factors); *Polaroid*, 287 F.2d at 495 (eight factors); *Malarkey-Taylor Assocs. v. Cellular Telecomms. Indus. Ass'n*, 929 F. Supp. 473, 477 (D.D.C. 1996) (citing *Polaroid*, 287 F.2d at 495) (eight factors), *with Interpace Corp. v. Lapp, Inc.*, 721 F.2d 460, 463 (3d Cir. 1983) (citing *Scott Paper Co. v. Scott's Liquid Gold, Inc.*, 589 F.2d 1225, 1229 (3d Cir. 1978)) (ten factors); *In re E.I. DuPont de Nemours & Co.*, 476 F.2d 1357, 1361 (C.C.P.A. 1973) (thirteen factors), *Pebble Beach Co. v. Tour 18 I Ltd.*, 155 F.3d 526, 541 (5th Cir. 1998) (seven factors); *Pizzeria Uno Corp. v. Temple*, 747 F.2d 1522, 1527 (4th Cir. 1984) (citing *Sun-Fun Prods., Inc. v. Suntan Research & Dev., Inc.*, 656 F.2d 186, 189 (5th Cir. 1981)) (seven factors); *Helene Curtis Indus., Inc. v. Church & Dwight Co.*, 560 F.2d 1325, 1330 (7th Cir. 1977) (citing *Carl Zeiss Stiftung v. Veb Carl Zeiss Jena, Steelmasters, Inc.*, 433 F.2d 686, 705 (2d Cir. 1970)) (seven factors); *SquirtCo v. Seven-Up Co.*, 628 F.2d 1086, 1091 (8th Cir. 1980) (six factors); *Team Tires Plus, Ltd. v. Tires Plus, Inc.*, 394 F.3d 831, 833 (10th Cir. 2005) (citing *King of the Mountain Sports, Inc. v. Chrysler Corp.*, 185 F.3d 1084, 1089 (10th Cir. 1999)) (six factors, with one factor combining two commonly separate criteria); *Conagra, Inc. v. Singleton*, 743 F.2d 1508, 1514 (11th Cir. 1984) (citing *Jellibean, Inc. v. Skating Clubs of Ga., Inc.*, 716 F.2d 833, 840 (11th Cir. 1983)) (seven factors).

44. *See* 4 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, § 24:30 (“[E]ach of the 13 federal circuit courts of appeal has developed its own version of the list and each appears to be jealous of its own formulation of factors.”).

45. *See infra* Table 1.

46. *See infra* Table 1.

47. *See infra* Table 1.

48. *See infra* Table 1.

federal circuits.

B. Research in Consumer Surveys by Legal and Marketing Scholars

Scholars explore consumer surveys from a variety of perspectives. In addition to the case-based analysis commonly seen in law reviews,⁴⁹ there is also a focus on consumer surveys in the marketing literature. Legal developments impact marketing strategies and, in turn, innovations in marketing research influence how courts apply legal protections.⁵⁰ Trademarks are also understood as strategic marketing tools.⁵¹ The result is

49. *E.g.*, Sandra Edelman, *Failure to Conduct a Survey in Trademark Infringement Cases: A Critique of the Adverse Inference*, 90 TRADEMARK REP. 746, 768–69 (2000) (analyzing how courts respond to a plaintiff’s failure to present survey evidence); Natalie-Claire Woods, *Survey, Survey Evidence in Lanham Act Violations*, 15 TRINITY L. REV. 67, 77–79 (2008) (comparing survey methods).

50. *See* Graeme W. Austin, *Trademarks and the Burdened Imagination*, 69 BROOK. L. REV. 827, 836 (2004) (“[M]arketing professionals still seek to produce the kind of unthinking responses to brands that much trademark law assumes occurs”); Ross D. Petty, *The Codevelopment of Trademark Law and the Concept of Brand Marketing in the United States Before 1946*, 31 J. MACROMARKETING 85, 90 (2011) (noting that judicial enforcement encouraged manufacturers to invest in brand marketing); Swann, *supra* note 2, at 945 (claiming that “explosive developments” in marketing since the 1960s translated into “more predictable, accurate and consumer-beneficial” results in trademark disputes); David S. Welkowitz, *Famous Marks Under the TDRA*, 99 TRADEMARK REP. 983, 994 (2009) (arguing that the expansion of infringement protection to include non-competing products solidified the role of trademarks as a marketing tool); Erin Lenhardt, Note, *Why So Glum? Toward a Fair Balance of Competitive Interests in Direct-to-Consumer Advertising and the Well-Being of the Mentally Ill Consumers it Targets*, 15 HEALTH MATRIX 165, 181 (2005) (stating that from a marketing perspective, trademark law protects brand equity, which is built through widespread exposure of the mark or brand name).

51. *See* Dorothy Cohen, *Trademark Strategy*, 50 J. MARKETING 61, 61 (1986) (arguing that marketers will have to become more familiar with trademark law as they become more active in setting trademark strategy); Dorothy Cohen, *Trademark Strategy Revisited*, 55 J. MARKETING 46, 46 (1991) (arguing that legal developments require marketers to become familiar with trademark law); Dennis S. Corgill, *Measuring the Gains of Trademark Infringement*, 65 FORDHAM L. REV. 1909, 1939 (1997) (arguing that a trademark can be a valuable marketing tool when it conveys information about the product “more efficiently” than other marketing devices); Alan S. Gutterman, *A Legal Due Diligence Framework for Inbound Transfers of Foreign Technology Rights*, 24 INT’L LAW. 976, 987 (1990) (claiming that trademarks can be valuable marketing tools because they can imply a certain level of quality); Angela L. Patterson, Comment, *With Liberty and Domain Names for All: Restructuring Domain Name Dispute Resolution Policies*, 40 SAN DIEGO L. REV. 375, 415 (2003) (arguing that trademarks are marketing tools because they can facilitate a customer’s ability to find and purchase items on a firm’s website); Ross D. Petty, *Naming Names: Trademark Strategy and Beyond: Part One—Selecting a Brand Name*, 15 J. BRAND MGMT. 190, 190 (2008) (examining the interplay between trademark law and marketing strategy in selecting a brand name and presenting strategies to limit the ability of rivals to select and use other brand names); John D. Shakow, Note, *Just Steal It: Political Sloganeering and the Rights of Trademark Holders*, 14 J.L. & POL. 199, 204 (1998) (stating that well known

a fairly rich body of information about consumer surveys from a welcome non-legal perspective.

For example, Stewart evaluated the characteristics of surveys that were introduced into evidence during an advertising misrepresentation case brought by the Federal Trade Commission against Kraft in the early 1990s.⁵² An article by Ford explored the impact of an important United States Supreme Court decision, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,⁵³ on the underlying rules of whether a survey is admissible as evidence at a trial.⁵⁴ A paper by Hastak, Mazis, and Morris examined the use of surveys in the development and process of policymaking.⁵⁵

Studies have also examined the role of consumer surveys in a variety of trademark contexts that are distinct from infringement cases. For example, Morrin, Lee, and Allenby measured the extent of harm from trademark dilution by using a Bayesian version of the associative network model.⁵⁶ The study found that even a single exposure to diluting brand stimuli reduced brand recall by about one-third on average.⁵⁷ The authors also used their findings to advise litigants on how to evaluate confusion factors in trademark dilution lawsuits.⁵⁸ Another article examined the reception of consumer surveys by courts in trademark dilution cases and found that courts remain skeptical of survey evidence in a variety of contexts.⁵⁹ In addition to dilution, scholars have reviewed court cases to evaluate what evidence courts consider in making findings of genericide, the cancellation of a trademark when consumers use the mark to describe a

trademarks are valuable marketing tools because they “enhance a company’s goodwill by creating positive associations with customers”).

52. David Stewart, *Deception, Materiality, and Survey Research: Some Lessons from Kraft*, 14 J. PUB. POL’Y & MARKETING 15, 21–23 (1995).

53. 509 U.S. 579 (1993).

54. Gary T. Ford, *The Impact of the Daubert Decision on Survey Research Used in Litigation*, 24 J. PUB. POL’Y & MARKETING 234, 234 (2005).

55. Manoj Hastak et al., *The Role of Consumer Surveys in Public Policy Decision Making*, 20 J. PUB. POL’Y & MARKETING 170, 170 (2001).

56. Maureen Morrin et al., *Determinants of Trademark Dilution*, 33 J. CONSUMER RES. 248, 249 (2006). The associative network memory model in psychology views memory as a network of nodes, each representing a piece of stored information and connected by links which reflect the strength of the association between two pieces of information. This theory casts the brand and the trademark as two nodes in memory. JOHN R. ANDERSON, *THE ARCHITECTURE OF COGNITION* (1983); Thomas K. Srull & Robert S. Wyer, Jr., *Person Memory and Judgment*, 96 PSYCH. REV. 59 (1989).

57. Morrin et al., *supra* note 56, at 253.

58. *Id.* at 254.

59. Bird, *supra* note 20, at 102; *see also* John P. Liefeld, *How Surveys Overestimate the Likelihood of Consumer Confusion*, 93 TRADEMARK REP. 939, 941 (2003) (arguing that surveys can overestimate likelihood of confusion by creating a response bias toward well-known companies).

generic category of products rather than a specific brand.⁶⁰

Development of guidelines and standards for survey evidence is also an important focus in the legal literature.⁶¹ Early research focused on increasing the likelihood that surveys would even be admissible as evidence in court.⁶² Later, Morgan proposed guidelines for developing survey research for use in court that would reflect explicit and implicit judicial perspectives of survey research methodology.⁶³ Morgan based his recommendations on a review of federal cases that incorporated survey research.⁶⁴ Legal treatises present elaborate guidelines for devising and assessing a survey's effectiveness.⁶⁵ Guidelines exist on survey design, population definition and sampling, data entry methods, interview techniques, and survey question structure in order to elicit accurate and unbiased results.⁶⁶ In general, these guidelines conform to standards in marketing theory and practice.⁶⁷

C. *Judicial Reception to Survey Research*

In judicial opinions, courts have directly discussed the importance of survey research in likelihood of confusion cases. Courts have called consumer surveys some of the most direct and persuasive evidence available to establish trademark infringement.⁶⁸ The absence of a consumer

60. Taylor & Walsh, *supra* note 23, at 160.

61. E.g., Larry C. Jones, *Developing and Using Survey Evidence in Trademark Litigation*, 19 MEMPHIS ST. U. L. REV. 471, 478–80 (1989) (discussing standards of when to conduct surveys and forms of survey questions); John Paul Reiner, *The Universe and Sample: How Good is Good Enough?*, 73 TRADEMARK REP. 366, 366–75 (1983) (discussing guidelines for the surveys).

62. Fred W. Morgan, *The Admissibility of Consumer Surveys as Legal Evidence in Courts*, 43 J. MARKETING 33, 38–39 (1979) (arguing that the increased use of surveys in court will enhance judges' sophistication in evaluating survey methodology); Sorensen & Sorensen, *supra* note 12, at 133 (arguing in 1955 that survey methods were sufficiently rigorous and reliable to be used in court).

63. Fred W. Morgan, *Judicial Standards for Survey Research: An Update and Guidelines*, 54 J. PUB. POL'Y & MARKETING 59, 68–69 (1990).

64. *Id.* at 59.

65. See, e.g., Benoît Gauthier, CIRCUM NETWORK INC., *Assessing Survey Research: A Principled Approach* (Mar. 2, 2003), http://www.circum.qc.ca/textes/assessing_survey_research.pdf (proposing guidelines for assessing a survey's effectiveness).

66. Shari Seidman Diamond, *Reference Guide on Survey Research*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 229, 236–69 (Joe Cecil & Dean Miletich eds., 2000).

67. See generally, DAVID A. AAKER ET AL., *MARKETING RESEARCH* (10th ed. 2009) (discussing marketing research theories and methodologies).

68. E.g., *Checkpoint Sys., Inc. v. Check Point Software Techs., Inc.*, 269 F.3d 270, 283 n.10 (3d Cir. 2001) (citing *Charles Jacquin Et Cie, Inc. v. Destileria Serralles, Inc.*, 921 F.2d 467, 476 (3d Cir. 1990)); *Vision Sports, Inc. v. Melville Corp.*, 888 F.2d 609, 615 (9th Cir.

survey has strongly suggested to some courts that the contents of the survey would have been unfavorable if administered.⁶⁹ One judge contrasted the “enormous sums” the plaintiff invested in market research for its product launch with its failure to commission a consumer survey in litigation, and thus drew an adverse inference from the absence of a survey.⁷⁰ A defendant, who has no burden of proof in trademark infringement cases, was even faulted for failing to produce a survey in court.⁷¹ Courts have expressed little sympathy even for small companies lacking the financial resources to shoulder the extraordinary costs of survey design and implementation.⁷²

Some commentators have similarly elevated consumer surveys to high evidentiary importance. A consumer survey is believed to attract so much attention that it can distract the court from other important evidence.⁷³ One commentator canvassing court cases remarked that when a plaintiff chooses not to present a consumer survey, the owner of the trademark may be perceived as “less than deadly serious about its case.”⁷⁴ Another stated that “[s]electing and performing a consumer survey for use as evidence in a trademark dispute represents one of the most important decisions made by

1989) (“An expert survey of purchasers can provide the most persuasive evidence of secondary meaning.”) (citing *Levi Strauss & Co. v. Blue Bell, Inc.*, 778 F.2d 1352, 1358 (9th Cir., 1985)); *McNeil Nutritionals, L.L.C. v. Heartland Sweeteners L.L.C.*, 566 F. Supp. 2d 378, 392 (E.D. Pa. 2008) (“[C]onsumer surveys are useful, and indeed the most direct method of demonstrating secondary meaning and likelihood of confusion . . .”) (quoting *Charles Jacquin Et Cie*, 921 F.2d at 476)).

69. See *Charles Jacquin Et Cie*, 921 F.2d at 475–76 (noting that while consumer surveys are not necessary to establish likelihood of confusion under the Lanham Act, a plaintiff’s failure to conduct a survey when it has the means to do so could lead a jury to believe that the results of the survey would be unfavorable to the plaintiff); *E.S. Originals Inc. v. Stride Rite Corp.*, 656 F. Supp. 484, 490 (S.D.N.Y. 1987) (finding it significant that the defendant failed to conduct a consumer survey when it had the means to do so, indicating that the defendant did not believe that it could demonstrate likelihood of confusion). But see Edelman, *supra* note 49, at 768–69 (concluding that some courts draw an adverse inference from a party’s failure to present survey data, while others treat a lack of survey data neutrally).

70. *Pharmacia Corp. v. Alcon Labs., Inc.*, 201 F. Supp. 2d 335, 373 (D.N.J. 2002).

71. *Gimix, Inc. v. JS&A Group, Inc.*, No. 80 C 6592, 1982 WL 52164, at *1006 (N.D. Ill. Jan. 13, 1982).

72. *Sports Traveler, Inc. v. Advance Magazine Publishers, Inc.*, 25 F. Supp. 2d 154, 164 (S.D.N.Y. 1998) (explaining that a plaintiff’s burden to produce probative evidence of confusion, such as by means of a consumer survey, is not conditioned on plaintiff’s financial resources).

73. See James Swire, Remarks at Panel Discussion, *Legal Standards for Consumer Survey Research*, 23 J. ADVERTISING RES. 19 19, 23 (1983) (describing how a defective survey harmed a litigant’s otherwise strong case).

74. 6 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, § 32:195.

trial counsel”⁷⁵ Still another called surveys “all but indispensable” in successfully demonstrating that a trademark is worthy of protection.⁷⁶

At the same time, judges have not been reluctant to discount consumer surveys if they conclude that the survey methodology was flawed. Consumer surveys in court undergo an aggressive examination for methodological flaws by experts hired by the opposing party.⁷⁷ As a result, criticisms related to virtually every aspect of a survey’s design—including selection of sample, interviewer bias, suggestive wording, and inappropriate purchasing conditions—often give judges reason to discount surveys before them.⁷⁸ The judge has to sift through the evidence to determine whether the criticisms are valid. This process places consumer surveys in the difficult position of being both supposedly essential evidence and highly vulnerable to criticism. This complicated role of surveys makes the need to test them empirically and understand their impact in trademark cases all the more compelling.

D. *Data-Driven Research on Consumer Survey Use in Federal Court*

Studies that look beyond individual cases to examine a substantial dataset of trademark case law are most relevant to our research. While case-based analysis has existed for decades, data-driven research is a more recent vintage. For example, Jacoby and Morrin reviewed several federal trademark infringement cases from 1994 through 1997 that were published in a commercial legal database.⁷⁹ The authors concluded, among other findings, that courts often heavily discount survey evidence and that a lack of convergence exists on how to measure the likelihood of confusion in trademark infringement cases.⁸⁰

More recent work systematically analyzes larger datasets. For example, one study examined the relationship between plaintiffs’ presentation of survey and actual confusion evidence and injunction outcomes in 126 federal trademark infringement cases between 2001 and

75. Robert H. Thornburg, *Trademark Surveys: Development of Computer-Based Survey Methods*, 4 J. MARSHALL REV. INTELL. PROP. L. 91, 91 (2004).

76. Joshua M. Dalton & Ilisa Horowitz, *Funny When You Think About It: Double Entendres and Trademark Protectibility*, 88 J. PAT. & TRADEMARK OFF. SOC’Y 649, 652 (2006).

77. See 6 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, § 32:178 (“It is notoriously easy for one survey expert to appear to tear apart the methodology of a survey taken by another.”).

78. *Id.* at § 32:171 (observing that courts discredit the evidential weight of deficient surveys).

79. Jacoby & Morrin, *supra* note 23, at 17.

80. *Id.* at 103–04.

2006.⁸¹ The authors found that plaintiffs received a modest improvement in outcomes through the presentation of actual confusion or survey evidence.⁸² The authors reported that their results are not statistically significant. This means that the authors could not sufficiently discount the reasonable possibility that the results could have occurred by chance.⁸³ Nonetheless, the authors noted that the results were still directionally interesting and showed that the odds of winning improved only slightly with presentation of actual confusion or survey evidence.⁸⁴

Of greatest relevance to our research is a thoughtful study by Barton Beebe.⁸⁵ While our focus here is the treatment of consumer surveys in federal courts, Beebe's work studied the varying application across circuits of multifactor tests for likelihood of confusion. Beebe examined 331 trial-level opinions that articulated the multifactor test of consumer confusion in trademark cases between 2000 and 2004.⁸⁶ He used regression trees to investigate what factors were most influential in determining likelihood of confusion in trademark infringement cases.⁸⁷

His results were consistent with the proposition that courts do not consider all factors in their circuit's multifactor test for likelihood of confusion.⁸⁸ Instead, Beebe found that courts simply examined a few

81. Sarel & Marmorstein, *supra* note 20, at 1417, 1423.

82. *Id.* at 1430.

83. By contrast, findings of statistical significance can make such a conclusion about empirical results. See, e.g., Lisa Faigman, *The "M Word" Symposium: An Interdisciplinary Adventure*, 23 HASTINGS WOMEN'S L.J. 3, 25 n.96 (2012) ("In scientific studies, if an observed difference does not rise to the level of 'statistical significance,' it essentially means that the observed findings could have occurred by chance."); Bradley W. Joondeph, *The Partisan Dimensions of Federal Preemption in the United States Courts of Appeals*, 2011 UTAH L. REV. 223, 237 (2011) ("[A] finding of statistical significance means that if the null hypothesis is true, there is less than a 5% chance that we would see the observed difference."). Expressed more technically, a non-statistically significant finding means that a statistical comparison (such as a t, F, or chi value statistic) did not reach the critical level of having a probability of appearing of .05 or less if all differences were purely random. This critical level is often denoted $p < .05$. Perry A. Zirkel, *Case Law for Functional Behavior Assessments and Behavior Intervention Plans: An Empirical Analysis*, 35 SEATTLE U. L. REV. 175, 201 n.160 (2011). As Zirkel explains:

Chi-square is a statistical procedure to determine whether the frequency counts in two or more categories in a sample . . . are differently distributed to a significant extent, i.e., that the frequency counts are not due to chance but are instead generally applicable with a high degree of probability to a population

Id. at 200 n.157.

84. Sarel & Marmorstein, *supra* note 20, at 1430.

85. Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CAL. L. REV. 1581 (2006).

86. *Id.* at 1593.

87. *Id.* at 1600.

88. *Id.* at 1599.

factors through “fast and frugal” heuristics.⁸⁹ In particular, Beebe discovered that a court’s finding regarding the “[s]imilarity of the marks,” “[p]roximity of the goods,” and “[s]trength of plaintiff’s mark,” favored confusion in a dominant share of plaintiff verdicts.⁹⁰ A subsequent study expanded on Beebe’s inquiry to look at 224 cases from the Southern District of New York that were published during a fifteen-year period.⁹¹ While Beebe’s primary focus was on the variation of application of the multifactor test across the federal circuits, Beebe’s analysis only briefly addressed survey data.⁹²

For our work on consumer surveys, the relevant question is what impact surveys have on the outcome of court cases. Our research builds on the aforementioned work on multifactor tests of trademark infringement and consumer surveys and offers a novel contribution to the literature on likelihood of confusion cases. The study methods and presentation of the data follow.

II. STUDY METHODS AND PRESENTATION OF DATA

A. *Data Collection*

The data for this study were collected from published opinions in trademark infringement cases involving confusion written by federal trial court judges in the United States between 2000 and 2006.⁹³ For the years 2000 through 2004, we obtained Beebe’s dataset. This is the same dataset that Beebe used in his examination of the multifactor tests of trademark infringement.

In addition to Beebe’s dataset, we hand-coded two additional years of data (2005 and 2006), following the same coding methods Beebe used for the prior years. We obtained data by gathering opinions using Westlaw and LexisNexis, two commercial legal research databases, and used search terms to capture any opinion which discussed a multifactor test for proving

89. *Id.* at 1602.

90. *Id.* at 1611, Table 4.

91. Kevin Blum et al., *Consistency of Confusion? A Fifteen-Year Revisiting of Barton Beebe’s Empirical Analysis of Multifactor Tests for Trademark Infringement*, 2010 STAN. TECH. L. REV. 3, 3–4 (2010).

92. Beebe, *supra* note 85, at 1622, 1641–42.

93. This paper focuses primarily on trademark infringement claims that sought to show consumer confusion between the senior (established) mark and the junior adoption of its mark. We do not explicitly explore disputes involving misappropriation, licensing, dilution, or false advertising, although we certainly hold open the possibility that our findings here have some relevance to these classes of cases.

trademark infringement in any meaningful way.⁹⁴

Similar to Beebe, we reviewed the judicial opinions from these results and eliminated any case that did not make substantial use of a multifactor test of the factors in Table 1. We also removed from the dataset any cases that involved counterfeiting, franchising, licensing, or a distribution agreement that might skew the results.⁹⁵ Like Beebe, we did not incorporate appellate decisions into our dataset to avoid redundancy.⁹⁶ We also did not consider claims of trademark dilution.

Our data collection and coding returned 202 additional cases. Adding these cases to Beebe's dataset of 311 created a total of 533 federal trademark cases for our entire dataset. The vast majority of these additions were cases published in 2005 and 2006. A small number of the cases added (fifteen) were opinions written between 2000 and 2004 that fit Beebe's criteria but were not included in Beebe's dataset, apparently because of delayed reporting by LexisNexis and Westlaw.

From this dataset, we examined how courts treated all likelihood of confusion factors relevant in their circuit. As noted earlier, each federal circuit has its own likelihood of confusion standard.⁹⁷ Coding whether a court concluded that a given factor favored or disfavored confusion was relatively straightforward. Courts typically state explicitly whether a given factor favored confusion, disfavored confusion, or had no effect. When a court did not discuss a factor, deemed it not relevant, or otherwise left it not clearly resolved, we coded accordingly.

Our dataset also notes whether or not the plaintiff or defendant submitted a survey. Our dataset reports whether a submitted survey was credited by the court. We considered a survey to be credited if the judge indicated that, despite any flaws, the survey contained at least some probative value in favor of the party submitting the survey. For example, if the plaintiff submitted a survey and the court found that the survey helped show a likelihood of confusion, despite problems with the survey's methodology, we considered that survey to be credited by the court.

Two caveats from our data collection process bear mentioning. First,

94. For the district courts of each circuit, Beebe ran the following search in Westlaw: ("trademark infringement" & confus!) or ("trademark mark" & "likelihood of confusion") & da(aft 12/31/1999 & bef 01/01/2005). For LexisNexis, Beebe ran the following search: ("trademark infringement" & confus!) or ((trademark mark) & "likelihood of confusion") and date(geq (01/01/2000) and leq (12/31/2005)). We followed these search criteria for the additional two years coded.

95. In these cases, the likelihood of confusion is often quite clear and the factors overwhelmingly favor the plaintiff.

96. Beebe, *supra* note 85, at 1593 (collecting only district court case data); Sarel & Marmorstein, *supra* note 20, at 1422–23 (describing the dataset).

97. *See infra* Table 1.

the data only include cases that actually went to court. To the extent that cases that settle differ from those that reach court, our findings do not generalize to the universe of infringement lawsuits. These caveats apply to similar studies on surveys as well.⁹⁸ Indeed, there may be a difference between those cases that reach court and those that do not. In those cases that reach court, both sides believe that they have a chance to succeed. Thus, our dataset will tend to include litigants who have the resources to litigate as well as the confidence that their side will win in court.

Second, although we believe that using LexisNexis and Westlaw to gather cases represents a substantially comprehensive review of the cases during the study period, the electronic databases do not report every judicial opinion in the United States. LexisNexis and Westlaw, although generally quite thorough, have their own methods of selecting opinions for inclusion in their database and exclude according to certain criteria.⁹⁹ In spite of this limitation, we have no reason to believe that the cases reported by Lexis and Westlaw are inaccurate reflections of federal trademark infringement court decisions nationwide.

B. *Plan of Analysis and Justification*

We performed our analysis in multiple stages. Our first task was to answer the basic question of how often litigants submitted consumer surveys as evidence in trademark infringement cases. In order to answer this question, we simply counted the number in which surveys were submitted relative to the number of total cases in our dataset.

Even this simple count data can reveal meaningful information about survey usage and court conduct. Most obviously, it reveals how often survey data is used in trademark infringement cases. This frequency can imply some information about influence of surveys over disputes. If the frequency of survey data is high, this supports the belief that survey evidence is indeed a common and even necessary proof to show likelihood of confusion.¹⁰⁰

The frequency of survey use also adds to the discussion on the

98. See Beebe, *supra* note 85, at 1597 (describing limitations of the data); Blum et al., *supra* note 91, at 4 (describing methodology for data analysis); Sarel & Marmorstein, *supra* note 20, at 1422–23 (explaining survey methodology).

99. See, e.g., Kimberly D. Krawiec & Kathryn Zeiler, *Common-Law Disclosure Duties and the Sin of Omission: Testing the Meta-Theories*, 91 VA. L. REV. 1795, 1832 & n.100 (2005) (noting the selection biases inherent when using Westlaw and LexisNexis).

100. E.g., Edelman, *supra* note 49, at 747 (“[S]urvey evidence has become de rigeur in trademark infringement cases.”); Simonson, *supra* note 16, at 364 (“[S]urveys are now routinely employed to prove likelihood of confusion, and a failure to introduce a survey into evidence often leads to harsh criticism by the courts.”).

influence of science on judicial decision making. An important discussion in the literature involves the susceptibility of judges to “junk science,” in part because of their lack of empirical and mathematical training.¹⁰¹ Combining this information with the frequency of favorable results for the survey user, we can speculate whether the mere presence of scientific evidence such as a consumer survey, regardless of survey quality, might generate a halo effect on the quality and persuasiveness of other evidence.

In addition to counting survey use, we classified our surveys according to which litigant submitted the evidence to the court. This variable segments the data according to whether a plaintiff or defendant submitted the evidence. In rare cases, both plaintiff and defendant submitted a survey in the same case, and we code such instances accordingly.

We separate plaintiff from defendant because the parties have different motivations in the production of evidence in a trademark dispute involving consumer confusion. Their differing motivations might influence the quality and emphasis of surveys that they construct. For example, the plaintiff holds the burden of persuasion and therefore would be more motivated to produce survey evidence to show consumer confusion.¹⁰² The defendant, by contrast, must merely show that the plaintiff’s evidence was insufficient.¹⁰³ Thus, the defendant may be less motivated to submit a survey that matches the complexity and expense of the plaintiff’s, if it chooses to commission one at all. Therefore, the quantity and impact of survey production might differ between plaintiff and defendant.

The second stage of our analysis, a series of logistic regressions, produces insight into the effectiveness of surveys. We avoid simply counting and cross-classifying survey usage with other variables. Such an analysis might be subject to errors in interpretation because it does not control for other evidence in the case.¹⁰⁴ Our logistic regressions of the

101. See, e.g., Michael I. Meyerson & William Meyerson, *Significant Statistics: The Unwitting Policy Making of Mathematically Ignorant Judges*, 37 PEPP. L. REV. 771, 776 (2010) (“[W]hile striving to avoid accepting ‘junk science’ into evidence, too many judges have permitted statisticians and others to allow ‘junk law’ into the courts.”); Paul S. Miller et al., *Daubert and the Need for Judicial Scientific Literacy*, 77 JUDICATURE 254 (1994) (arguing for systematic judicial education on science and the scientific method).

102. See Irina D. Manta, *In Search of Validity: A New Model for the Content and Procedural Treatment of Trademark Infringement Surveys*, 24 CARDOZO ARTS & ENT. L.J. 1027, 1062 (2007) (describing the opposite evidentiary burdens of plaintiff and defendant); see also Ruth M. Corbin & Arthur Renaud, *What’s Good for the Goose Is Bad for the Gander: Why Confusion Surveys for Plaintiff and Defendant Should Be Different*, 16 INTELL. PROP. J. 177 (2003) (arguing that confusion surveys for defendants and plaintiffs should not be similar).

103. Manta, *supra* note 102, at 1062 (noting lower burden of persuasion for defendants).

104. See Matthew Sag et al., *Ideology and Exceptionalism in Intellectual Property: An*

form $\Pr(\text{Confusion Found}) = f(\text{Predictive Variables})$ enable such control variables to be accounted for.

For additional variables, we first inquire into how a court treats survey evidence. Not all surveys are treated equally by judges. Some courts reject survey evidence out of hand as being insufficiently persuasive or lacking scientific rigor; other courts are more receptive. We account for this differential treatment by evaluating whether a survey was credited by a court. As noted earlier, we consider a survey to be credited if the judge indicated that, despite any flaws inherent in the survey, it still contained at least some probative value.

Given that survey evidence can be a potentially powerful indicator of consumer confusion in the minds of some judges,¹⁰⁵ we posit that the submission of survey evidence that the court believes has some probative value will influence the outcome of the case. As noted earlier, plaintiffs and defendants have different evidentiary burdens in trademark cases. Thus, we also test the impact of plaintiff-credited surveys and defendant-credited surveys as separate variables.

We also consider whether the strength of a plaintiff's claim on grounds unrelated to survey evidence impacts the influence of consumer surveys in court. For example, if the plaintiff already has a strong case of infringement, a consumer survey might not be especially influential. A weak case of infringement, by contrast, might be materially aided by a survey. The impact of survey evidence might also be influenced by whether the plaintiff or defendant submits it in a weak or strong case context. For example, if the defendant submits a survey when the plaintiff presents strong non-survey evidence of infringement, the survey might be treated differently than when a plaintiff's non-survey case is weak.

In addition, the impact of a credited or non-credited survey in evidence might be influenced by the strength of non-survey evidence in the case. For example, a plaintiff's chances of winning might be irreparably damaged by a non-credited survey if its non-survey evidence is of middling

Empirical Study, 97 CAL. L. REV. 801, 829 n.142 (2009) ("A control variable is a variable that is held constant in order to analyze the relationship between other variables without interference.").

105. *E.g.*, *Tone Bros., Inc. v. Sysco Corp.*, 28 F.3d 1192, 1204 (Fed. Cir. 1994) (describing the persuasiveness of survey evidence (citing *Co-Rect Prods. v. Marvy! Adver. Photography, Inc.*, 780 F.2d 1324, 1333 n.9. (8th Cir. 1985))); *Miller's Ale House, Inc. v. Boynton Carolina Ale House, LLC*, No. 09-80918-CIV, 2009 WL 6812111, at *17 (S.D. Fla. Oct. 13, 2009) (noting that consumer survey evidence is the best evidence of confusion (citing *Investacorp, Inc. v. Arabian Inv. Banking Corp.*, 931 F.2d 1519 (11th Cir. 1991))); *Jewish Sephardic Yellow Pages, Ltd. v. DAG Media, Inc.*, 478 F. Supp. 2d 340, 364-65 (E.D.N.Y. 2007) (noting that consumer surveys are "particularly helpful in divining . . . the principal significance' of the term to the consuming public" (quoting *Horizon Mills Corp. v. QVC, Inc.*, 161 F. Supp. 2d, 208, 216-17 (S.D.N.Y. 2001))).

strength. Alternatively, a plaintiff marshaling strong non-survey evidence might be forgiven if the survey evidence proves flawed or incomplete.

To measure the impact of non-survey evidence on survey receptivity, we incorporate as variables the court's evaluation of certain *Polaroid* factors in the case. Incorporating every factor, however, remains a problem because not all circuits apply all *Polaroid* factors.¹⁰⁶ To ensure equanimity across federal circuits, we limited our consideration only to factors that all federal circuits rely upon when determining likelihood of infringement. Out of the twelve *Polaroid* factors that courts consider, only four are considered by all thirteen federal circuits.¹⁰⁷ We eliminated the "evidence of actual confusion" factor as a variable because courts sometimes consider consumer surveys to be evidence of actual confusion, thus making reliance on this factor as non-survey evidence redundant.¹⁰⁸

That left three factors that were uniformly considered across the federal circuits: the similarity of the litigants' marks to one another, the proximity of litigants' goods to one another in the marketplace, and the strength of the plaintiff's mark amongst the public. We incorporate these variables into our regression analysis to show the impact of non-survey proof on consumer survey evidence.

Although only three variables of the possible twelve are considered in this study, these variables can be helpful in illuminating our results. Beebe found that these three variables were disproportionately predictive of the outcome of cases.¹⁰⁹ Beebe hypothesizes a "stampede effect" when courts review these initial factors and then "stampede" the remaining factors in order to conform to the initial factor outcomes.¹¹⁰ We thus conclude that relying on these factors is potentially telling about the influence of the

106. See *infra* Table 1.

107. See *infra* Table 1.

108. See, e.g., 6 J. THOMAS MCCARTHY, in MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, § 32:184 ("Several courts, when assembling the evidence within a likelihood of confusion framework of factors such as the *Polaroid* Eight, have put survey evidence under the heading of 'actual confusion.'"). McCarthy is also careful to conclude, however, that survey evidence is circumstantial and should not be treated as direct evidence of a likelihood of confusion. According to McCarthy, direct evidence can only come from communications from customers or the exceedingly rare testimony of an individual confessing that they were once a confused consumer. *Id.*

109. Beebe, *supra* note 85, at 1622.

110. *Id.* at 1582, 1617–22. This occurs in spite of judicial admonitions that all factors must be thoroughly considered in likelihood of confusion cases. See, e.g., *Arrow Fastener Co. v. Stanley Works*, 59 F.3d 384, 400 (2d Cir. 1995) ("[I]t is incumbent upon the district judge to engage in a deliberate review of each factor, and, if a factor is inapplicable to a case, to explain why The steady application of *Polaroid* is critical to the proper development of trademark law"). The court cautions, however, that courts need not slavishly recite all *Polaroid* factors in each and every case. *Id.* (quoting *Orient Express Trading Co., Ltd. v. Federated Dep't Stores, Inc.*, 842 F.2d 650, 654 (2d Cir. 1988)).

Polaroid factors on survey evidence.

III. STUDY FINDINGS

In this section, we report the results of our analyses. We also discuss how our findings contribute to our understanding of survey evidence and its influence. While some results reinforce prevailing perceptions about consumer surveys, others are unexpected and surprisingly illuminating.

A. *Frequency of Use of Consumer Survey Evidence*

Our data show that survey use is not widespread in reported trademark infringement litigation. Out of the 533 cases reviewed for our dataset, only eighty-nine (16.6%) discuss survey evidence. A plaintiff submitted a survey in seventy-four of these cases, defendants in twenty-three, and in eight cases both plaintiff and defendant submitted a survey.¹¹¹

Thus, even though survey evidence has attracted significant scholarly and judicial attention, we find no evidence that surveys are used by a majority or even a large plurality of litigants to prove likelihood of confusion in federal court. Our results seem to contradict the exhortation by some courts and commentators that consumer survey evidence is necessary or even strongly recommended to prove trademark infringement.

The apparent implication is that consumer surveys are not especially useful in likelihood of confusion cases. If a company is going to court without survey evidence, and presumably believing themselves to have a reasonable chance of success without such evidence, then a firm may want to consider seriously the time and expense involved before commissioning survey evidence in a trademark dispute. Global conclusions cannot be made from this result alone, but this surprising result appears to throw into question the notion that survey evidence is commonplace and an essential component of trademark infringement cases.

Even though our results show that survey evidence is not present in the majority of federal cases, one should take care not to dismiss the importance of survey evidence entirely. In spite of these results, a real possibility exists that consumer surveys still have an indirect effect on likelihood of confusion cases even if they do not appear in judicial opinions.

For example, surveys can be conducted internally by a company and never be used in court. A firm may commission a survey to determine

111. Beebe reported survey usage in twenty percent of the cases in his dataset. Beebe, *supra* note 85, at 1641.

whether a trademark infringement suit is viable against a rival. The survey may produce unfavorable results and get discarded, resulting in the lawsuit never taking place. It is thus possible that the weakest survey evidence never leaves the corporation for which it was commissioned.

If a company commissions a survey and its results are encouraging, it can still have an impact outside of the federal courtroom. Firms, most likely large firms with significant resources, may use survey results as leverage to extract a favorable settlement from an infringing rival. The rival may evaluate the survey results, realize that its chances of successfully defending against a trademark infringement claim are small, and settle the claim. Surveys can be and are instrumental in settlement negotiations.¹¹² Thus, the mere potential of survey evidence used against a potential litigant could compel it to resolve the dispute out of court.

Even if litigation begins, we cannot divine from published court opinions how often surveys are used, not in litigation, but outside of the courtroom. It is therefore possible that we underestimate the use of surveys in trademark litigation since the data we analyze exclude those cases in which surveys were used but never surfaced in a court's judicial opinion. Although we cannot say for certain, what we observe in the federal court system may merely be the "tip of the iceberg" of survey usage in trademark-related disputes.¹¹³ Nevertheless, the absence of consumer surveys in the vast majority of court decisions is striking and suggests that observers who view surveys as essential may overstate their influence.

Another possible explanation of ours is that survey evidence may have been submitted, but never written about, in the judge's opinion. Such conduct would result in an artificially low percentage of survey cases reported. To account for this potential problem, we randomly selected legal briefs from thirty non-survey cases to see if they mentioned survey evidence or the submission of a survey by the plaintiff or defendant. None were found. This suggests that judicial opinions ignoring submitted surveys are unlikely and that consumer survey evidence is indeed sparse in our data.

B. Ability to Help Prove Likelihood of Confusion in Trademark Infringement Cases

Although we now know that consumer surveys are typically not used in reported trademark infringement cases, it remains to be seen whether

112. See Jacob Jacoby, *Experimental Design and the Selection of Controls in Trademark and Deceptive Advertising Surveys*, 92 TRADEMARK REP. 890, 890 (2002) (noting that surveys are used for both litigation and settlement).

113. *Id.*

surveys have any impact in proving or disproving a likelihood of confusion. We begin our analysis with a simple count of wins and losses in our dataset. Out of the 533 decisions mentioned above, there were 60 instances when the court could not determine whether the plaintiff proved a likelihood of confusion because the case had factual issues that remained unresolved. Of the 473 remaining cases, courts found likelihood of confusion 242 times (approximately 51.1% of cases). Courts failed to find likelihood of confusion 231 times (approximately 48.8% of cases).¹¹⁴ In short, plaintiffs win about half of the time in trademark infringement cases without accounting for the presence of survey evidence.

We further our analysis through the estimation of five logistic regression models, displayed in Table 2. The table consists of five models, with one model listed per column. Each model tests the impact of different survey-related variables on the court's finding of a likelihood of confusion. Model 1 is the simplest of the five. Model 1 examines the relationship between probability of a court finding a likelihood of confusion and whether or not the plaintiff submitted a survey (logit). This model simply inquires whether the plaintiff's submission of survey, regardless of other factors such as the quality of the survey or the strength of non-survey evidence, impacts on a court's likelihood of confusion finding. This model also does not consider the impact of defendant-submitted evidence. Our results reveal that no statistically significant relationship exists between plaintiffs submitting a survey and a finding of likelihood of confusion. While this may appear counterintuitive at first glance, and appears to offer further proof of the relative irrelevance of survey evidence, we need to withhold judgment until accounting for further variables.

Model 2 adds an additional variable by examining the correlation between defendant-submitted surveys and likelihood of confusion outcomes. Presumably, because a defendant needs only to counter the plaintiff's evidence in order to win, defendants proffer fewer surveys than plaintiffs. Our dataset included twenty-three cases where the defendant submitted survey evidence, and a court reached a decision in twenty of those cases.

Although Model 2 raises issues of small sample size,¹¹⁵ it also yields an interesting result: while a plaintiff's submission of a survey does not

114. This roughly corresponds to Sarel and Marmorstein's finding of a fifty percent success rate in their dataset. Sarel & Marmorstein, *supra* note 20, at 1424.

115. See Matthew Laroche, *Is the New York State Court of Appeals Still "Friendless?"*: An Empirical Study of Amicus Curiae Participation, 72 ALB. L. REV. 701, 754 (2009) ("Where sample sizes are small, it is appropriate to exercise caution when interpreting the results."); see also 6 J. THOMAS MCCARTHY, IN MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, *supra* note 6, at § 32:171 (noting that small sample size problematic for survey design).

have a significant relationship on a finding of confusion, a defendant's submission negatively correlated with a finding of a likelihood of confusion. This is not surprising in itself, but is intriguing when considered in tandem with the lack of influence from plaintiffs' submissions. This implies that courts treat plaintiff surveys and defendant surveys differently.

A number of reasons may explain this phenomenon. First, plaintiff-submitted survey evidence may be perceived as unsurprising by some courts, since the plaintiff's burden of proof motivates it to submit as much evidence as possible. When a defendant submits survey evidence, by contrast, it is a rare and thus an unexpected event. A defendant-submitted survey may signal to the court that the defendant is serious about defending its trademark usage. The survey may also signal to the court that the defendant is confident about winning its case, although if a defendant were so sure that consumers would not be confused then a survey would not be necessary. A defendant-submitted survey may neutralize the scientific "halo effect" that would accompany a plaintiff-submitted survey not rebutted by other seemingly scientific evidence. Finally, a defendant-submitted survey merely needs to show that one cannot conclusively establish consumer confusion, an easier burden than the plaintiff must bear.

Model 3 introduces yet another variable, the impact of whether or not a consumer survey was credited by the court. Model 3 investigates both the impact of a plaintiff-submitted survey that is credited and a defendant-submitted survey that is credited. Once again, the impact of a credited survey differs according to which litigant submitted the survey. Our results reveal that when the plaintiff submits a survey and the court credits that survey as having probative value favorable to plaintiff, that credited survey is strongly and significantly correlated with a finding of likelihood of confusion by the court.¹¹⁶

The finding of Model 3 is not unexpected. A consumer survey represents powerful scientific evidence of likelihood of confusion by soliciting the opinions of the very group of consumers that is impacted by the presence of competing marks. Such data should help plaintiffs prove

116. The probability that this finding in our data was the result of chance is miniscule, represented by a p-value of less than .0001. As one author succinctly explains:

The "p" value is a statistical measure of probability. For example, a p value of less than .05 indicates that the statistical likelihood that the observed result occurred by chance is less than 5%, $p < .01$ means less than 1%, and so forth. A lower p value indicates a higher statistical significance.

Hon. Donald E. Shelton et al., *An Indirect-Effects Model of Mediated Adjudication: The CSI Myth, The Tech Effect, and Metropolitan Jurors' Expectations for Scientific Evidence*, 12 VAND. J. ENT. & TECH. L. 1, 18 n.60 (2009) (citing Mark J. Schervish, *P Values: What They Are and What They Are Not*, 50 AMER. STATISTICIAN 203 (1996)).

their case in court. Our results reveal that it is likely that plaintiff-submitted survey evidence, when credited by the court, increases the probability of a likelihood of confusion finding.

Model 3 for defendants produces a different result. Our results report that a court-credited survey submitted by the defendant does not have a statistically significant impact on the outcome of the dispute. Apparently it is the presence of the defendant-submitted survey, rather than its methodology, quality, or persuasiveness that has the favorable impact on the court. This does not necessarily mean that defendants who are conducting surveys should ignore basic survey methodology, but it may imply that defendant-submitted surveys are not subjected to the same rigor as their plaintiff-submitted counterparts. The fact that the defendant took the time and effort to submit a survey appears to be enough. Submitting a survey that is good enough to actually be credited by a court does not appear to further strengthen a defendant's case.

In Model 4, we test a set of different but important variables: the *Polaroid* factors that underlie a finding of likelihood of confusion. As noted earlier, the similarity of the litigants' marks to one another, the proximity of litigants' goods to one another in the marketplace, and the strength of the plaintiff's mark amongst the public are the three *Polaroid* factors that are most suitable for testing. These factors are shared by all federal circuits and appear to have a strong influence over likelihood of confusion findings. We test whether a court's determination that a given factor favors or disfavors finding likelihood of confusion actually has an impact on the outcome of the dispute. Our results find that these three factors are all influential. Each of these three factors, when found to favor a finding of likelihood of confusion by the court, has a statistically significant outcome on the result of the case. For example, when a court finds that the trademarks are similar enough to favor a likelihood of confusion finding, the court is significantly more likely to find likelihood of confusion overall. This finding also works in reverse. When a court concludes that one of these favors militates against a likelihood of confusion finding, the court is overall less likely to reach an overall likelihood of confusing finding. These three factors impact the ultimate outcome of trademark infringement disputes.

Model 5, our most complex model, considers the impact of consumer surveys in the context of the strength of the litigants' non-survey evidence. In the previous models, we discerned whether the presence of a survey, the presence of a credited survey, or the author of the survey had a significant impact on the likelihood of confusion finding. These models are incomplete since the possibility exists that the non-survey *Polaroid* factors may overwhelm the survey evidence and render it immaterial. Model 5

addresses that problem by considering whether these variables still hold influence when taking into account the impact of the *Polaroid* factors on judicial outcomes.

This model reveals that survey evidence indeed remains influential in likelihood of confusion cases, even when the influence of *Polaroid* factors is separately taken into account. The pattern of significance corresponds to that in Model 3. This verifies the inferences made above with respect to plaintiffs' and defendants' credited and non-credited surveys. A credited plaintiff survey increases the probability of a likelihood of confusion finding, a non-credited plaintiff survey decreases it, and any defendant survey increases the probability of a no likelihood of confusion finding. Even when the non-survey likelihood of confusion factors are taken into account, survey evidence remains influential.

C. *A Cost-Benefit Analysis of Consumer Surveys*

The five models reveal intriguing and useful information about the influence of survey evidence in court. In addition, the algebraic form of the logistic regression—upon which these models are based—allows us to quantify the precise benefit of consumer survey evidence in a variety of contexts. Not only can we learn about the directional impact of survey evidence, but we can also understand more about the magnitude of that impact in court.

Table 3 presents the magnitude of survey impact. Each row reflects one possible combination of the three dominant likelihood of confusion factors in a trademark infringement dispute. Courts could reach three different possible conclusions regarding each factor: the factor favors a confusion finding, the factor favors a no-confusion finding, or the factor favors neither result. The number of permutations for three likelihood of confusion factors, each having three possible evidentiary outcomes, totals twenty-seven. Accordingly, Table 3 presents twenty-seven rows, each representing one possible permutation.

The first three columns in Table 3 reflect these permutations. The next four columns present the probability of confusion given the combinations of factors in the first three columns in various instances.¹¹⁷ Column 4 presents the probability of a likelihood of confusion finding when no survey is presented. Columns 5, 6, and 7 present the probability of a likelihood of confusion finding in cases where only the plaintiff submitted a survey, only the defendant submitted a survey, and when both

117. These probabilities are all based on Model 5 in Table 2 and assume that the surveys are credited.

parties submit a survey, respectively.

The last two columns measure the impact of plaintiff-submitted and defendant-submitted surveys in each scenario.¹¹⁸ Each number represents the change in probability of likelihood of confusion as a result of the survey, given the particular scenario. For example, reviewing the first row shows that when the court finds that all three *Polaroid* factors favor confusion, a plaintiff-submitted survey increases the chances of a confusion finding by 3.2%. A defendant-submitted survey reduces the chances of a confusion finding by one percent under similar conditions.

As Table 3 shows, the impact of a survey varies widely depending on the surrounding evidence. In some situations, survey evidence can have a profound impact. For example, consider the criteria in Row 8 of Table 3. This row assumes that the trademarks are similar (favoring confusion), the products at issue are not similar (disfavoring confusion), and that it is uncertain whether the strength of the mark favors confusion. In this case, submission of a credited consumer survey would increase the probability of a confusion finding from .097 to .864, assuming the defendant does not submit a survey. This represents an apparent 76.7% increase in the probability that a likelihood of confusion finding will occur from the submission of a credited consumer survey by the plaintiff. In the same scenario, the defendant's submission of a credited survey appears to reduce the probability of a likelihood of confusion finding by 61.5%.

Although showing mathematical impacts such as these can appear overly convincing, one should interpret Table 3 cognizant of its limitations. As noted earlier, only a fraction of our dataset contains cases where survey evidence is evaluated by the court. Segmenting this portion of the dataset by the three *Polaroid* factors only slices the size of the dataset into smaller portions. The result of such segmentation is that the probabilities and impacts in Table 3 should be interpreted as showing trends, rather than a precise measurement of future effects.

These trends, however, reveal that surveys play varying roles for plaintiffs. Surveys seem to be most helpful to plaintiffs when non-survey proof is of middling strength. In each of the five scenarios where plaintiff-submitted surveys appear most influential, there is a mix of "confusion," "no confusion," and "neither" findings.¹¹⁹ Conversely, plaintiff-submitted

118. Each scenario assumes that when a plaintiff submits a survey, the defendant does not submit survey evidence. The reverse applies to the column showing the impact of a defendant-submitted survey. Indeed, the empirical likelihood of both plaintiff and defendant submitting a survey in trademark infringement litigation is quite low.

119. For the convenience of the reader, the five evidentiary conditions when plaintiff's survey is most influential are reproduced here from Table 3:

surveys do not appear to be influential when the plaintiff's case is particularly weak or strong. In each of the five conditions where plaintiff-submitted survey evidence is least influential, two or more of the factors indicate "confusion" or "no confusion."¹²⁰ In fact, the weakest and strongest possible evidentiary conditions, whereby all three factors either favor or do not favor confusion, are scenarios where plaintiff-submitted survey evidence appears to have little effect. When key non-survey evidence is especially strong, a plaintiff appears not to need survey evidence in order to show a likelihood of confusion. When key non-survey evidence is especially weak, survey evidence will typically be insufficient to overcome a finding of no likelihood of confusion.

In the case of defendant-submitted evidence, similar principles apply. Defendant-submitted surveys appear to be most influential in cases involving middling evidence. Indeed, four of the five conditions in which surveys are most influential for the plaintiff are the same conditions that are most influential for the defendant. The conditions in which defendant-submitted surveys are least influential also roughly correspond to plaintiff-submitted surveys. Thus, defendant-submitted surveys play a comparable

SIMILARITY OF MARKS	SIMILARITY OF PRODUCTS	STRENGTH OF PLAINTIFF'S MARK	IMPACT OF PLAINTIFF'S SURVEY
Confusion	No Confusion	Neither	0.767
Confusion	Neither	No Confusion	0.756
No Confusion	Confusion	Confusion	0.756
Neither	Confusion	No Confusion	0.753
Neither	Neither	Neither	0.739

Out of the five conditions, the no confusion/confusion/confusion scenario appears out of place because it appears to represent a strong evidentiary case on behalf of the plaintiff. However, the similarity of marks prong is overwhelmingly influential in trademark cases. An adverse finding on this prong alone may be so devastating that it needs survey evidence to bolster it, even though the other two dominant factors favor confusion. *See Beebe, supra* note 85, at 1623 (concluding that "[t]he data clearly show that the similarity of the marks factor is by far the most important factor in the multifactor test.").

120. For the convenience of the reader, the five evidentiary conditions when plaintiff's survey is least influential are reproduced here from Table 3:

SIMILARITY OF MARKS	SIMILARITY OF PRODUCTS	STRENGTH OF PLAINTIFF'S MARK	IMPACT OF PLAINTIFF'S SURVEY
Neither	No Confusion	No Confusion	0.057
Confusion	Confusion	Confusion	0.032
No Confusion	Neither	No Confusion	0.032
No Confusion	No Confusion	Neither	0.018
No Confusion	No Confusion	No Confusion	0.002

role in trademark infringement cases. Where key *Polaroid* factors are indeterminate, defendant-submitted surveys can play an influential role. Where key *Polaroid* factors strongly point in one direction, defendant-submitted surveys have little impact.

While remaining cognizant of the limitations mentioned above, these probability measurements can enable primitive calculations to determine whether a survey is financially viable under a given set of conditions. For example, assume that a plaintiff-submitted survey improves the probability of success by 10.5% or .105. A relatively small increase in the probability of success might imply that a plaintiff should not submit a survey in this instance.

This depends, however, on the expected benefit of the survey relative to the value of the potential favorable verdict. Assume that the plaintiff assesses the expected value of its claim, net of attorney's fees and costs, at \$1,000,000. If a survey costs \$100,000 to administer,¹²¹ and the survey produces an incremental increase in probability of .105, the litigant should conduct the survey because the benefit of the survey ($\$1,000,000 * 1.05 = \$105,000$) is greater than the cost (\$100,000) of creating it.

Thus, using Table 3, a litigant could improve his or her decision-making about whether to conduct a consumer survey by performing a realistic assessment of its likelihood of confusion claim and financial data based on court behavior in prior similar cases. Such calculations can be made for any combination of the likelihood of confusion factors provided based upon an objective assessment of the strength of one's own legal claim. Such a calculation is far from an exact science, given the limitations of the data and the fact-sensitive nature of each trademark infringement claim, but using this information may improve decisions and reduce uncertainty in outcomes.

IV. REEVALUATING THE ROLE OF SURVEY EVIDENCE

We briefly consider, in light of these data, whether the role of survey evidence in trademark infringement cases needs to be redefined. Courts should refrain from making an adverse inference when litigants do not submit a consumer survey as an evidentiary part of their case. As discussed earlier, a number of judges have concluded that the absence of a survey implies a weak case or outright laxness by the parties. However, as our findings show, the significant majority of litigants do not use consumer

121. See Peter Weiss, *The Use of Survey Evidence in Trademark Litigation: Science, Art or Confidence Game?*, 80 TRADEMARK REP. 71, 85 n.30 (1990) ("A fairly reliable mini-survey can be done for \$5,000 or less. The cost of a major survey tends to start at around \$30,000 and can run up to \$100,000 or more.").

surveys in reported likelihood of confusion cases.

Thus, it is doubtful that the significant majority of trademark litigants, and their associated attorneys, are lax. Instead, the choice is likely a calculated decision to not submit such evidence based upon time, cost, and administrative burdens. As noted earlier, consumer surveys are expensive to design and implement.¹²² Plaintiffs with limited resources will not be able to afford one. Litigants are also deterred from submitting survey evidence because courts can be excessively picky about consumer survey design,¹²³ making it difficult to submit survey evidence that will reliably pass judicial muster. There are also, as our findings show, calculated financial reasons for choosing not to submit a survey given the assessed strength of the claim.

The Trademark Trial and Appeal Board, the entity responsible for hearing specialized trademark cases such as appeals from patent examiners and opposition proceedings filed against trademark applicants, explicitly does not require parties to submit surveys and refuses to draw negative inferences from the lack of survey evidence.¹²⁴ The Board remarked that “[n]either party is obligated, in a proceeding before the Board, to spend the effort and expense to obtain such evidence.”¹²⁵ The federal circuits reviewing likelihood of confusion cases should follow the Board’s lead by clearly and explicitly refusing to make an adverse inference from a lack of survey evidence.

In the alternative, however, we acknowledge that there may be situations in which a court may feel compelled to make a negative inference in the absence of such evidence. If a negative inference must be made from the lack of a survey, it may be appropriate under conditions cited by Edelman.¹²⁶ Edelman writes that the lack of survey evidence should hurt a plaintiff’s case only if the case is suitable for a survey, there is ample time to conduct a survey, no logistical obstacles are present, and there is no other persuasive evidence of actual confusion.¹²⁷ These criteria, perhaps with the additional consideration of financial ability, create a situation that could possibly create a negative inference. However, due to the difficulties in survey construction and submission, such an adverse inference should be rare indeed.

Although courts should not fault litigants for failing to marshal survey

122. Bird, *supra* note 20, at 104.

123. See Bird, *supra* note 13, at 283–84 (discussing *Stuart Hall Co. v. Ampad Corp.*, 31 U.S.P.Q. 2d 1468, 1471 (W.D. Mo. 1994)).

124. *McDonald’s Corp v. McClain*, 37 U.S.P.Q. 2d 1274, 1276–77 (T.T.A.B. 1995).

125. *Id.* at 1277.

126. Edelman, *supra* note 49, at 766–67.

127. *Id.*

evidence, such evidence can still retain significant persuasive power under certain circumstances. As a result, courts should evaluate the methodological validity of survey evidence according to well established criteria. There are already a number of sources, mentioned earlier, that offer detailed and intelligent guidelines on all aspects of survey design.¹²⁸ Our results show, though indirectly, that survey evidence is vulnerable to arbitrary factors. For example, courts appear to treat plaintiff-produced surveys and defendant-produced surveys differently. The mere submission of a survey by a defendant appears to help her case, while a plaintiff-submitted survey can potentially hurt her case if the court deems it flawed. Although the burdens of proof are different between plaintiffs and defendants, the evidentiary weight of a survey's design and implementation should not vary significantly simply due to which party submitted it.

Our research also adds a new dimension to the literature on survey evidence—the status of the party making the submission—a characteristic rarely studied in survey research.¹²⁹ In theory, courts should treat survey evidence equally, regardless of which party, plaintiff or defendant, submitted the evidence. Our study finds, however, that courts treat survey evidence differently according to whether the submitting litigant is a plaintiff or a defendant. Future research studying the impact of consumer surveys on court decisions or the construction of survey evidence in litigation should incorporate this dimension. This research opens up the possibility that survey evidence submitted in trademark dilution,¹³⁰ deceptive advertising,¹³¹ and genericide¹³² cases may also be susceptible to litigant-status preferences by the court.

Finally, the findings in this paper reinforce calls to reform the current state of the Lanham Act multifactor test used in likelihood of confusion cases. The current likelihood of confusion criteria potentially misleads litigants into believing that all of the factors will be given equal weight. While some courts affirm that all factors are equally important and must be considered, a number of courts do not give each factor similar importance.¹³³ Our findings, as well as Beebe's findings,¹³⁴ that three

128. AAKER ET AL., *supra* note 67; Diamond, *supra* note 66; Gauthier, *supra* note 65.

129. Ruth M. Corbin & Arthur Renaud, When Confusion Surveys Collide: Poor Designs or Good Science?, 94 Trademark Rep. 781, 781–82 (2004) (discussing the lack of survey research that takes account of party status).

130. Bird, *supra* note 20, at 102.

131. Stewart, *supra* note 52, at 15–16.

132. Taylor & Walsh, *supra* note 23, at 161.

133. Frank Mead, Note, *Cocaine, Coffee Mugs, Sex, and Bug Killing Floor Wax: Welcome to the Realm of Parody and the Likelihood of Confusion*, 21 T. JEFFERSON L. REV. 305, 308–09 (1999) (“The test applied by the courts is called the ‘likelihood of confusion test.’ However, the only thing that is confusing is the application of this test. To begin,

factors are disproportionately influential, only reinforce the uncertainty of the criteria. As Table 3 shows, even without survey evidence, when the three aforementioned factors (similarity of the marks, proximity of the goods, and strength of the plaintiff's mark) favor confusion, courts find in favor of the plaintiff more than 95% of the time. Other factors, such as the sophistication of consumers and the similarity of advertising methods, are less substantial considerations than the aforementioned critical factors.

Our findings reinforce the results of another study finding that the consumer sophistication, a factor in determining likelihood of confusion, was actually influenced by the similarity of brand names.¹³⁵ This conclusion supports our finding that similarity of marks is an influential factor in determining likelihood of confusion. Our findings also generally encourage further work that studies the impact of individual factors in the likelihood of confusion criteria.¹³⁶

Two possible recommendations can emerge from the issue of disproportionate influence of the three dominant *Polaroid* factors. The first recommendation is to establish a nationwide standard of likelihood of confusion criteria across all federal circuits. Currently, each federal circuit uses a different multifactor test, and despite multiple recent Supreme Court opinions on trademark law, the Court has yet to adopt a nationwide standard. Such a standard would help to develop more uniform precedent and increase certainty of outcomes. The second recommendation is to reduce the likelihood of confusion determination to the three dominant factors mentioned earlier plus a fourth factor, actual evidence of confusion, which would allow for admission of consumer surveys and other direct evidence.

A less radical change would be to retain the non-dominant criteria, but enable courts to use them only when the first dominant four factors do not resolve the question of confusion. Given the explanatory power already found, such situations would be infrequent, but would still allow judicial flexibility if the court deemed it appropriate. In spite of the apparent need for greater uniformity and clarity for likelihood of confusion criteria, the Supreme Court has had numerous opportunities to establish uniform factors when it decided trademark infringement cases and has refused to do so.

each circuit court applies the test differently, utilizing only those factors it deems important.”) (footnotes omitted).

134. Beebe, *supra* note 85, at 1591.

135. Daniel J. Howard et al., *The Effects of Brand Name Similarity on Brand Source Confusion: Implications for Trademark Infringement*, 19 J. PUB. POL'Y & MARKETING 250, 261–62 (2000).

136. Thomas R. Lee et al., *Trademarks, Consumer Psychology, and the Sophisticated Consumer*, 57 EMORY L.J. 575, 578–79 (2008); Jerre B. Swann, *Sophistication and the Sciences*, 97 TRADEMARK REP. 1309, 1314–15 (2007).

Thus, the chance of such a reform remains unlikely.

V. LIMITATIONS AND FURTHER RESEARCH

The results reviewed and the cost-benefit tools built in the previous sections are only as valid as the data used to obtain them. We only examined a seven-year period (2000–2006). It is not clear whether our results are stable over time. Also, economic fluctuation may impact whether litigation is brought (litigation is costly) and whether a survey is conducted (surveys are expensive too), but this would only exacerbate the low incidence of survey usage. In addition, we do not know about surveys that were conducted, found to have unfavorable results, and never introduced into evidence.

We do not know the role that surveys played in litigation that was settled and never reached trial. It is possible, if not likely, that surveys play a very different role in cases that settle than in those that go to trial. To the extent that surveys force settlement, our estimate of the degree to which they are used could be vastly understated. The limitation of the size of the defendant-submitted survey dataset is also a limitation. Defendants frequently limit themselves to critiquing the plaintiff's survey methodology and results. As such, our conclusions regarding defendants can only be considered tentative owing to small samples.

Like any scientific research, no survey is perfect. As one survey expert explained, "I know of no survey, in trademark litigation or elsewhere, which fulfills one hundred percent the requirements of its model."¹³⁷ As a result, future research is needed to examine under what conditions surveys are most persuasive to a court. Some, although by no means all, of the limitations in this research can be addressed in future work.

Of particular interest for future research are cases where a litigant submitted a survey that was credited by the court but the litigant still failed to persuade the court overall. For plaintiffs, this would occur when a plaintiff submitted a survey showing consumer confusion, the court considered the survey credible, and the plaintiff still failed to prove infringement. For defendants, this would occur when a defendant submitted a survey showing little or no confusion, the court considered the survey credible, and the plaintiff proved a likelihood of confusion anyway. These cases may offer insight into what conditions negate the impact of an otherwise persuasive survey. Such information might be useful for litigants

137. Robert C. Sorensen, *Survey Research Execution in Trademark Litigation: Does Practice Make Perfection?*, 73 TRADEMARK REP. 349, 351 (1983).

facing an opponent who has presented well-constructed consumer surveys in court. Another intriguing topic for further study would be to examine which factors evaluated to determine survey quality (i.e., universe, sample size, interview techniques) are most vulnerable to judicial critique in order to reveal issues most critical for future survey design. Although encouraging studies exist about the nature of consumer surveys, there is much to learn about their influence.

CONCLUSION

Brands are big business. They define the public face of organizations and have immense power to shape the perception of consumers toward a product or service. Brands are as vulnerable as they are valuable. Firms can tarnish their own brands through strategic errors. Upstart rivals can outflank brands with their own marketing strategies, making seemingly new and innovative products appear out of date by comparison. While the law does not interfere with fair competition, trademarks protect brands from inappropriate consumer confusion created by a rival brand so similar that it can be mistakenly associated with the established trademark.

The consumer survey is one of the most widely studied methods of proving consumer confusion in court. The consumer survey has a radiance of scientific certainty about it. Its scientific nature and direct questioning of consumers may seem to some as disproportionately persuasive to a judge lacking expert training. At first glance, one can wonder whether consumer surveys are perceived by some as the magic bullet of trademark litigation.

This study reveals that we now know more than we once did about consumer surveys' use and importance. Consumer surveys are not used as often as some would imply. They are also not universally influential. While surveys can prove persuasive under certain conditions, they might not be useful for litigants with particularly weak or strong evidence. Presentation of survey evidence still remains an individual decision left to the attorney and his trademark-owning client. However, litigants can now make decisions that are at least partially informed by prior judicial activity and experience. In some cases, the most efficient survey is the one that is never commissioned in the first place.

While this article sheds light on consumer surveys, there is still much to learn. Survey evidence exists in a variety of contexts outside trademark infringement litigation, and an intriguing issue is to what extent the conclusions offered here apply in other contexts. Empirical studies like this one and those completed by other authors offer a springboard from which to conduct further research. Although survey evidence may always remain in some sense a "black box" of litigation, the further study of such

2012] CONSUMER SURVEYS IN TRADEMARK INFRINGEMENT 1049

evidence can provide important knowledge about the construction, use, and reception of potentially highly persuasive and influential scientific evidence in court.

TABLE 1: FACTORS FEDERAL COURT CIRCUITS CONSIDER IN LIKELIHOOD OF CONFUSION CASES

FACTOR	CIRC.													FED*
	1	2	3	4	5	6	7	8	9	10	11	DC.		
1 Similarity of the marks	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2 Proximity of the goods	x	x	x	x	x	x	x	x	x	x	x	x	x	x
3 Evidence of actual confusion	x	x	x	x	x	x	x	x	x	x	x	x	x	x
4 Strength of plaintiff's mark	x	x	x	x	x	x	x	x	x	x	x	x	x	x
5 Defendant's intent	x	x	x	x	x	x	x	x	x	x	x	x	x	
6 Sophistication of consumers	x	x	x			x	x	x	x	x			x	x
7 Similarity of advertising and marketing methods	x		x	x	x	x	x		x	x	x			
8 Similarity of sales facilities	x			x	x		x					x		x
9 Likelihood of bridging the gap		x	x			x			x				x	
10 Comparative quality of the goods		x											x	
11 Length of time of concurrent use without evidence of actual confusion			x											x
12 Similarity in parties' sales efforts			x											

* We excluded from this study factors only used in the federal circuit.

Insufficient cases came from that circuit to produce meaningful results

Source: Barton Beebe, An Empirical Study of the Multifactor Tests for Trademark Infringement, 94 CALIF. L. REV. 1581 (2006).

2012] CONSUMER SURVEYS IN TRADEMARK INFRINGEMENT 1051

TABLE 2: BINARY LOGISTIC REGRESSIONS RELATING SURVEYS TO PROBABILITY OF FINDING LIKELIHOOD OF CONFUSION { $Pr(C)$ }

VARIABLE	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5
Plaintiff's Survey	-.166	-.112	-1.728 ^c		-2.634 ^d
Defendant's Survey		-.912 ^b	-1.207 ^a		-2.007 ^a
Plaintiff's Survey Credited			3.490 ^e		6.714 ^e
Defendant's Survey Credited			.744		-.845
Similarity of Marks Favoring Confusion				2.210 ^e	2.214 ^d
Similarity of Marks Favoring No Confusion				-3.156 ^e	-3.637 ^e
Similarity of Products Favoring Confusion				1.235 ^e	1.324 ^e
Similarity of Products Favoring No Confusion				-2.763 ^e	-3.037 ^e
Strength of Plaintiff's Mark Favoring Confusion				1.140 ^b	1.267 ^c
Strength of Plaintiff's Mark Favoring No Confusion				-2.104 ^e	-2.417 ^e
Constant	.068	.098	.091	-1.390 ^b	-1.411 ^b
-2 LOG LIKELIHOOD	655.09	651.47	619.09	198.48	176.39

a – $p < .10$ b – $p < .05$ c – $p < .01$ d – $p < .001$ e – $p < .0001$

TABLE 3: PROBABILITIES OF FINDING LIKELIHOOD OF CONFUSION {Pr(C)} FOR COMBINATIONS OF JUDICIAL FACTORS AND THE IMPACT OF SURVEYS

ROW NUMBER	SIMILARITY OF MARKS	SIMILARITY OF PRODUCTS	STRENGTH OF PLAINTIFF'S MARK	Pr(C No SURVEYS)	IMPACT OF PLAINTIFF SURVEY	IMPACT OF DEFENDANT SURVEY
1	Confusion	Confusion	Confusion	0.968	0.032	-0.010
2	Confusion	Confusion	Neither	0.893	0.104	-0.035
3	Confusion	Confusion	No Confusion	0.428	0.550	-0.280
4	Confusion	Neither	Confusion	0.888	0.110	-0.037
5	Confusion	Neither	Neither	0.691	0.302	-0.119
6	Confusion	Neither	No Confusion	0.166	0.756	-0.541
7	Confusion	No Confusion	Confusion	0.275	0.682	-0.417
8	Confusion	No Confusion	Neither	0.097	0.767	-0.615
9	Confusion	No Confusion	No Confusion	0.009	0.352	-0.332
10	Neither	Confusion	Confusion	0.765	0.230	-0.085
11	Neither	Confusion	Neither	0.478	0.504	-0.243
12	Neither	Confusion	No Confusion	0.076	0.753	-0.627
13	Neither	Neither	Confusion	0.464	0.517	-0.253
14	Neither	Neither	Neither	0.196	0.739	-0.505
15	Neither	Neither	No Confusion	0.021	0.541	-0.500
16	Neither	No Confusion	Confusion	0.040	0.671	-0.597
17	Neither	No Confusion	Neither	0.012	0.397	-0.374
18	Neither	No Confusion	No Confusion	0.001	0.057	-0.055
19	No Confusion	Confusion	Confusion	0.079	0.756	-0.626
20	No Confusion	Confusion	Neither	0.024	0.565	-0.519
21	No Confusion	Confusion	No Confusion	0.002	0.111	-0.106
22	No Confusion	Neither	Confusion	0.022	0.552	-0.508
23	No Confusion	Neither	Neither	0.006	0.269	-0.256
24	No Confusion	Neither	No Confusion	0.001	0.032	-0.031
25	No Confusion	No Confusion	Confusion	0.001	0.060	-0.057
26	No Confusion	No Confusion	Neither	0.000	0.018	-0.017
27	No Confusion	No Confusion	No Confusion	0.000	0.002	-0.002

2012] CONSUMER SURVEYS IN TRADEMARK INFRINGEMENT 1053

<i>ROW NUMBER</i>	<i>Pr(C PLAINTIFF SURVEY)</i>	<i>Pr(C DEFENDANT SURVEY)</i>	<i>Pr(C BOTH SURVEYS)</i>
1	0.999	0.609	0.989
2	0.998	0.305	0.963
3	0.978	0.038	0.698
4	0.998	0.293	0.961
5	0.992	0.104	0.873
6	0.922	0.010	0.381
7	0.957	0.019	0.540
8	0.864	0.006	0.249
9	0.361	0.000	0.029
10	0.995	0.145	0.910
11	0.982	0.046	0.739
12	0.829	0.004	0.202
13	0.981	0.043	0.728
14	0.935	0.013	0.430
15	0.563	0.001	0.063
16	0.711	0.002	0.114
17	0.409	0.001	0.035
18	0.058	0.000	0.003
19	0.835	0.004	0.209
20	0.588	0.001	0.069
21	0.113	0.000	0.007
22	0.574	0.001	0.066
23	0.275	0.000	0.019
24	0.033	0.000	0.002
25	0.061	0.000	0.003
26	0.018	0.000	0.001
27	0.002	0.000	0.000