
28. Empirical studies of trademark law *Barton* *Beebe**

Contents

- I. Introduction
- II. Studies of Trademark Registration
- III. Trademark Distinctiveness
- IV. The Likelihood of Confusion Test
 - A. Courts' Use of the Multifactor Test
 - B. The Likelihood of Confusion and Sponsorship Confusion
 - C. The Likelihood of Confusion in Context
 - D. The Role of Survey Evidence
- V. The Likelihood of Dilution Test
 - A. Anti-dilution Law in the Courts
 - B. Dilution Over Time
 - C. Experimental and Survey Evidence of Dilution
- VI. Conclusion References

I. INTRODUCTION

In one of the earliest efforts at bringing empirical methods to bear on trademark law, Judge Jerome Frank of the Second Circuit Court of Appeals questioned “some adolescent girls and their mothers and sisters, persons I have chosen at random,” about the facts of the case then before him in 1948. In the case, both the plaintiff and the defendant used the mark “Seventeen” for their respective products (*Triangle Publ'ns*, 167 F.2d 969 (1948)). Judge Frank reported: “I have been told uniformly by my questionees that no one could reasonably believe that any relation existed between plaintiff’s magazine and defendants’ girdles.” In the half-century that has passed since Judge Frank’s survey, we have undoubtedly come a long way in the sophistication and breadth of the empirical methods we employ in the adjudication of trademark disputes and in the study of trademark law. Indeed, the past two decades have seen an especially significant expansion in the empirical study of the workings of trademark law, ranging from the straightforward “case-counting” and systematic content analysis of reported court opinions to highly creative and innovative studies employing such varying resources as internet search

* John M. Desmarais Professor of Intellectual Property Law at the New York University School of Law.

engines, historic phonebooks, surveys of trademark lawyers, and massive datasets covering trademark registrations in the U.S. and Europe.

This chapter attempts a brief review of the present state of the empirical study of trademark law.¹ Section II covers studies of trademark registration. Section III considers studies of the concept of trademark distinctiveness. Section IV turns to empirical studies of the likelihood of confusion cause of action, while Section V covers empirical studies of trademark dilution.

II. STUDIES OF TRADEMARK REGISTRATION

In 2013, the Office of the Chief Economist of the U.S. Patent and Trademark Office (USPTO) released the USPTO *Trademark Case Files Dataset* (TCFD) (Graham, Hancock, Marco, and Myers, 2013), which the USPTO has since updated annually. The TCFD represents an extraordinary resource. As of 2015, it contains detailed information on all 7.6 million trademark applications filed at the USPTO from 1982–2015, including filing basis, prosecution history, ownership, assignment, goods and services classifications (the so-called “Nice classes” after the Nice Classification of Goods and Services), and the characteristics of the applied-for marks themselves. The dataset also contains partial information on a significant proportion of applications filed from 1870–1981. Graham, Hancock, Marco, and Myers (2013) provide an extensive overview of the data (then current through 2012) with a particular focus on trends over time. They show the substantial increase in the annual number of trademark applications filed at the USPTO from approximately 62 000 in 1982 to over 300 000 in 2011 (and the number has since increased to almost 370 000 in 2015). With Beebe (2011), they attribute the significant spike in trademark applications during 1999–2000 to the dot-com boom.² The authors further observe that with the introduction in November, 1989 of the intent-to-use filing basis, intent-to-use applications have come to make up a majority of trademark applications filed at the USPTO, though use-based applications still constitute a majority of the applications that advance to registration, owing in part to the significantly higher rates of abandonment before registration of intent-to-use applications.

In a forthcoming study, Beebe and Fromer (2018) analyze the TCFD to show severe levels of trademark depletion and congestion on the USPTO’s Principal Register. They define trademark depletion as the process by which an increasing proportion of potential trademarks are precluded from registration by already registered trademarks in one or more classes of goods or services. Trademark congestion is the process by which, for any particular trademark that has already been claimed, that mark is claimed by an increasing number of different trademark owners. Beebe and Fromer (2018) analyze the TCFD data in light of the most frequently used words and syllables in American English, the most frequently occurring surnames in the U.S., and all possible one-syllable words in English given spelling and pronunciation constraints. These data show that rates of word mark depletion and congestion have been increasing over the past three decades and have reached chronic levels, particularly in various important economic sectors. The data also show that trademark applicants are

¹ For a review of empirical studies of the economic role and function of trademarks, see Schautschick and Greenhalgh (2015).

² Gerhardt and McClanahan (2013) also note that this spike followed the USPTO’s introduction of electronic filing for trademark applications in October, 1998.

increasingly resorting to second-best marks that are more complex, as measured by character, syllable, and word count, and that take the form of neologisms rather than standard English words or frequently occurring surnames. One significant limitation of the TCFD is that it does not present data on the grounds for refusal of the trademark applications that the USPTO refused to approve for publication. To make up for this shortfall, Beebe and Fromer systematically downloaded from the USPTO website the full text of all trademark Office actions issued by the USPTO from 2003—the year when the USPTO began to post online its Office actions—through 2016. They report clear increases over this period in the proportion of applications receiving refusals on the ground that the applied-for mark is confusingly similar to an already registered mark, particularly in such economic sectors as apparel fashion and hightechnology goods. In light of these findings, Beebe and Fromer assert that the “ecology of the trademark system is breaking down, with mounting barriers to entry, increasing consumer search costs, and an eroding public domain.”

Sheff (2014) has also used the TCFD to analyze the impact of federal anti-dilution protection on registration practice at the USPTO. This work is considered below in the discussion of empirical studies of trademark dilution law.

Before the USPTO released the TCFD in 2013, Google agreed in 2010 to host on its servers the USPTO’s raw and somewhat haphazardly organized xml data that eventually formed the basis of the TCFD. Two early projects on trademark registration relied on this xml data. ‘Is the Trademark Office a Rubber Stamp?’ (Beebe, 2011) revealed overall and class-specific trademark application publication and registration rates over time at the USPTO (rates that the USPTO still does not publicly announce). The article established that the overall publication rates for both use-based and intent-to-use applications filed from 1989–2007 were the same, at .76, and that these rates varied little annually except for applications filed during the 1999–2000 period, when the publication rates declined to as low as .70 for use-based applications filed in 2000 and .68 for intent-to-use applications filed in that year. As for registration rates, the article noted the substantial difference between the annual registration rates for use-based applications, which closely tracked the annual publication rates for such applications, and the annual registration rates for intent-to-use applications, which averaged .37 from the inauguration of the intent-to-use basis in November, 1989 through 2007, and which were thus drastically lower than the annual publication rates for such applications. As noted above, this difference is largely due to many intent-to-use applicants’ failure to file statements of use after their applications had been published. Though the article is largely descriptive, it asks why trademark publication rates at the USPTO remain significantly higher than 50 percent. It proposes one possible explanation: that—particularly for use-based applications—if an applicant is in doubt as to whether its application will succeed, it may settle for protection of its mark as an unregistered mark rather than risk rejection by the USPTO. Combined with the cost and effort required for registration, the result is that some trademark owners which do not perceive a significantly better than even chance of success in registration may decide against applying.

Gerhardt and McClanahan (2013) also used the USPTO’s raw xml data as the basis of their article ‘Do Trademark Lawyers Matter?’, which analyzed the degree to which the involvement and quality of legal representation—rather than proceeding *pro se*—affected trademark applicant success rates at the USPTO. They focused on the 5.5 million trademark applications filed at the USPTO from 1984–2012. They chose a start date of 1984 because the data contained incomplete information on the presence of legal counsel for applications filed before this year. Their analysis of registration rates (rather than publication rates) covered

only the period from 1984–2010 because applications typically take 12–24 months to proceed through examination and potential opposition to registration or refusal. (This two-year buffer for the study of registration rates is standard practice in the literature.) They defined *pro se* applications as applications for which the “attorney-name” field in the data was left blank. As Gerhardt and McClanahan readily admit, this definition of *pro se* applicants is problematic. Among the most frequent such “*pro se*” filers in their data were large corporations such as Twentieth Century Fox Film (New Corp.) and Avon Products, which no doubt had access to sophisticated in-house legal counsel but filed their applications through paralegals or other administrators who left the attorney field blank. Nevertheless, the authors mitigate this problem to a great degree by taking into account the frequency with which certain attorneys and *pro se* filers filed trademark applications at the USPTO. Overall, they show that during the period studied, attorney-filed applications had a significantly higher publication rate (82 percent) than *pro se* applications (60 percent), particularly when applications met with an Office action from the USPTO (72 percent vs. 45 percent). Registration rates were also significantly different, with attorney-filed applications enjoying a 60 percent registration rate, far above the 42 percent registration rate for *pro se* applicants. Further emphasizing the importance of attorney involvement, Gerhardt and McClanahan (2013) identified 27 940 applicants that filed at least one *pro se* application and one attorney-filed application. For such applicants, publication and registration rates were significantly higher for applications filed by an attorney. Finally, the authors present compelling and detailed evidence that for both *pro se* and attorney-filed applications, increased experience levels of the filers correlated very strongly with increased publication and registration rates.

Working from an original dataset rather than the TCFD, Carpenter and Garner (2015) studied trademark applications that received an Office action refusing registration on the grounds that the applied-for trademark was “scandalous” or “immoral” under Section 2(a) of the Lanham Act, 15 U.S.C. § 1052(a). To do so, they developed a dataset consisting of 232 trademark applications filed from 2001–11 that received a Section 2(a) refusal by searching on the USPTO’s Trademark Electronic Search System for trademarks containing words likely to trigger such a refusal. Of these 232 applications, only two ultimately overcame the Section 2(a) refusal. Carpenter and Garner reported that 73 percent of the applications they studied resulted in refusals where the examiner did not discuss the context in which the mark was used, notwithstanding well-established doctrine requiring examiners to do so. Examiners used dictionary definitions as the *only* source of support for their refusals in 70 percent of the applications studied. Unsurprisingly, 97 percent of the applications studied were filed by individuals or small businesses, and 80 percent of the applicants abandoned their applications upon receiving a Section 2(a) refusal. Separate from their dataset of 232 applications, the authors additionally focused on five potentially scandalous terms appearing in applications from 2000–10 to show significant inconsistencies in USPTO review of these terms. Finally, they note that 53 percent of the applied-for marks were still in use in some fashion, regardless of whether they were ultimately refused registration, which suggests that “the morality bar is not effective at keeping immoral trademarks out of the marketplace” (Carpenter and Garner, 2015).

Turning to studies of trademark registration in the European context, Von Graevenitz, Greenhaigh, Helmers, and Schautschick (2012) provided a descriptive and econometric analysis of the phenomenon of “trademark cluttering,” which they defined as occurring when trademark registers contain “such a large number of unused or overly broad trademarks . . . that the cost of creating and registering new marks substantially increase for other applicants.”

Because European trademark law does not impose as stringent a use requirement on registrants as U.S. law does, and because cancellations for non-use are in practice very rare in the European system, trademark cluttering has emerged as a serious problem in Europe in recent years.³ Using trademark registration data from the Oxford Firm Level Intellectual Property database (Helmets, 2011) and financial data from Bureau van Dijk's Financial Analysis Made Easy, they developed a dataset consisting of 201 021 UK or European Community trademark registrations from 27 161 firms registered in the UK over the period 2000–07. They showed that over the period studied, firms tended to file more widely across more Nice classes, with medium-sized and large firms doing so to a greater extent than smaller firms. They used a linear regression model to show that growth in firms' trademark registration activity is largely independent of asset growth. Interestingly, they also identified in the data a "feedback mechanism" through which firms respond to increases in their competitors' trademark registration activity by increasing their own registration efforts—a process that the authors suggest may lead to increased cluttering. The authors additionally used data from the Office for Harmonization of the Internal Market (OHIM, now called the European Union Intellectual Property Office (EUIPO)) covering trademark applications at the office over the period 1996–2010. These data show, among other things, that after the UK Intellectual Property Office (UKIPO) ceased in 2007 to examine trademark applications for relative grounds for refusal (which considers whether the applied-for mark is confusingly similar to an already registered mark), UK firms' applications at the UKIPO faced an increased probability of being opposed relative to applications by firms from other countries; meanwhile, UK firms' applications at OHIM enjoyed a reduced probability of being opposed relative to applications by firms from other countries. The authors speculated that this may show UK firms' efforts before the 2007 policy change to use OHIM applications to circumvent UKIPO's relative grounds examination. Finally, the authors presented compelling evidence of simultaneous application activity in which firms apply for multiple different marks for one product even though they expect ultimately to use only one of the registered marks. This is particularly problematic in the pharmaceutical sector, where firms submit to public health regulators multiple invented names per drug in anticipation of the rejection of many of these proposed names out of a heightened concern for consumer confusion or misinterpretation of drug names. The result is a "surplus" of registered marks in the pharmaceutical sector in Europe that are unused and merely clutter the register.

Von Graevenitz (2013) extended his study of European firms' strategic trademark registration behavior in the face of public health regulators' invented pharmaceutical name review. He took advantage of the enlargement of the European Union (EU) in 2004, when ten new countries joined the EU, to estimate the proportion of surplus trademarks registered as Community trademarks. Each new accession country brought with it its own ability to veto invented drug names, which increased significantly the likelihood of rejection. Von Graevenitz modelled firm behavior to hypothesize that firms would increase after the 2004 EU enlargement the number of names they submitted to invented name review and for trademark registration. Using a dataset covering 597 450 trademark applications at OHIM from 1997–2009, he employed a difference-in-differences approach and a nearest neighbor matching approach to compare trademark application conduct by pharmaceutical firms before

³ In their highly influential review of European trademark law, Kur and Knaak (2011) cite a survey conducted by the *Institut für Demoskopie Allendach* suggesting that European trademark lawyers believe cluttering to be a problem.

and after the 2004 enlargement. He concluded that invented name review caused pharmaceutical firms to increase their trademark application filings from 10–15 percent per year, with the cost of developing and registering surplus invented names ranging from \$21 million to \$49 million per year. Von Graevenitz proposed that these costs could be minimized by reforms to the use requirement in European trademark law. More generally, he suggested that his findings may challenge the conventional view that trademark registers are not susceptible to depletion.

III. TRADEMARK DISTINCTIVENESS

In order to qualify for protection, a trademark must be “distinctive” of source. Distinctiveness is arguably the central requirement of trademark protection and among the most important concepts in trademark law. It is conventionally held that trademarks may be inherently distinctive of source or they may acquire source distinctiveness through use in commerce, advertising, or other means of informing consumers that the trademark is a source designation. As set forth most authoritatively (at least in the U.S.) in the Second Circuit opinion in *Abercrombie & Fitch v. Hunting World*, 537 F.2d 4 (1976), trademarks qualify as inherently distinctive of source when, due to their nature, consumers immediately perceive them as source indicators. In the case of word marks, words typically qualify as inherently distinctive if they are “fanciful” or coined (e.g., ACCENTURE for consulting services), “arbitrary” in the sense that they bear no semantic relation to the product (APPLE for high-technology goods), or merely “suggestive” of the product’s characteristics (IVORY for soap). Words that are descriptive of the product or its merits (AMERICAN for airline services) are traditionally understood to lack inherent distinctiveness. Their users must show that such descriptive marks have acquired distinctiveness over time in order to qualify for trademark protection. Generic terms (e.g., “Lite Beer” for low-calorie beer) are incapable of source distinctiveness. Together, this framework forms the “*Abercrombie* spectrum”, consisting of fanciful, arbitrary, and suggestive marks that are inherently distinctive, descriptive marks that must show acquired distinctiveness to gain protection, and generic marks that do not qualify for protection.

Through a series of clever experiments (which they described especially clearly and in great detail), Lee, DeRosia, and Christensen (2009) mounted a compelling challenge to the *Abercrombie* spectrum and in particular its understanding of how consumers perceive descriptive marks. Specifically, the authors showed that contextual indicators of meaning on product packaging such as the size, stylization, framing, and positioning of a word mark can instill the word mark with source distinctiveness even when the word is otherwise descriptive of the product—or even when the word is generic with respect to the product. In essence, while the conventional view has held that descriptive marks can only acquire secondary meaning over time through use and advertising, the authors showed that descriptive marks are capable of acquiring secondary meaning “*ab initio* through non-lexical cues” on product packaging. Such cues evoke the consumer’s “perceptual schema for product packages” and guide the consumer to the conclusion that the descriptive mark is being used as a brand name.

Lee, DeRosia, and Christensen (2009) conducted three studies. In the first, 210 individuals answered an online questionnaire about various images of product packaging for packaged cookies. This packaging featured one of various hypothetical trademarks from each of the five *Abercrombie* categories (fanciful, arbitrary, suggestive, descriptive, and generic). The packaging otherwise remained constant in appearance. The trademark appeared prominently

on the package in a context in which product brand names conventionally appear. The test respondents ultimately perceived the descriptive marks to be just as source distinctive as the suggestive marks—and indeed over one-quarter of the respondents perceived the generic mark to be source indicating. The second study significantly expanded on the first. In total, 930 individuals answered an online questionnaire about images of product packaging for four different categories of goods that carried 31 different hypothetical marks from the various *Abercrombie* categories. The authors found that “all of the descriptive marks were statistically equivalent to the suggestive marks in terms of source indication.” In their third study, the authors varied the number of contextual indicators on the product packaging that suggested that the word mark was being used as a brand name. The results strongly suggest that the contextual indicators are what instilled the descriptive marks in the previous studies with source distinctiveness.

Based on their findings, Lee, DeRosia, and Christensen (2009) urged that the descriptiveness bar to protection be removed in situations where the “placement, prominence, and style” of the mark will prompt consumers to recognize that it is being used as a source indicator. The authors recognized that “competitive need” offers an alternative justification for the denial of protection to descriptive marks that have not developed secondary meaning over time. But they urged a reformulation of this justification so that only “competitively essential” words would be denied protection.

Though the study provides an important and much-needed corrective to the orthodoxy of the *Abercrombie* spectrum, one limitation of the study is that its findings arguably do not apply to situations in which a mark is used independently of product packaging—for example, when the mark is used in aural communications, in product reviews, or in other media coverage of the mark.

In a different and highly creative approach to the study of the workings of trademark distinctiveness, Ouellette (2014) showed that Google search engine results can be used to estimate the strength of trademarks that are the subject of litigation (and, as discussed in the next section, Google results can also greatly aid in the assessment of the likelihood of confusion). Indeed, she argued that “Google does a better job than courts in evaluating the factual issue of consumer perceptions of trademarks.” Ouellette tracked federal court opinions over a one-year period from December 2011 through November 2012 that addressed trademark distinctiveness or the likelihood of consumer confusion. After filtering out opinions involving disputes between related companies, licensing disputes, and trade dress cases, she ended up with a dataset consisting of 88 cases. Within two weeks of the issuance of each opinion, Ouellette conducted Google searches, set to the location of the case venue (and in a later rerunning of the protocol, more broadly to the United States), for the relevant trademarks in the case, searching both for the trademarks alone as well as for the trademarks in combination with various keywords describing the field of use specified by the owner of the asserted mark. To assess the distinctiveness of the asserted mark, she analyzed the first page of organic, non-paid search results for that mark (and for the mark and keywords referring to its field of use) and counted the number of hits (out of a total of ten) that referred to the source identified by the asserted mark. Ouellette established a scheme by which the number of hits indicated that the mark was either unprotectable (0–2 hits), weak (3–5), medium (4–8), strong (8–10), or famous (10), and found that this scheme correlated with the court’s assessment of the strength of the asserted mark in 64 (or 73 percent) of the 88 cases she studied. Ouellette closely analyzed the cases in which Google and the courts disagreed and argued that, at least

on the issue of distinctiveness, the assessment yielded by a Google search rather than the court's assessment was correct in each instance.

IV. THE LIKELIHOOD OF CONFUSION TEST

The central question in trademark infringement analysis is whether the defendant's trademark as used on the defendant's goods is sufficiently similar to the plaintiff's trademark as used on the plaintiff's goods that consumers will be confused as to the true source of the defendant's goods. In most trademark systems around the world, judges assess this likelihood of consumer confusion as to source by recourse to a set of factors. This test typically considers such factors as the similarity of the plaintiff's and defendant's marks; the strength of the plaintiff's mark; the proximity of the parties' goods, sales, and marketing channels; the defendant's intent in adopting its mark; the sophistication of the relevant consumers; and any anecdotal or survey evidence of confusion among consumers.

A. Courts' Use of the Multifactor Test

To determine which factors drive the multifactor test for the likelihood of consumer confusion in U.S. federal courts, Beebe engaged in the systematic content analysis of all reported federal district courts opinions over the period 2000–04 that made substantial use of a multifactor test to assess the likelihood of consumer confusion (Beebe, 2006). Working from an original dataset of 331 opinions, Beebe further sought to show the degree to which judges engage in “coherence-based reasoning” (Simon, 2004) and “fast and frugal” heuristics (Gigerenzer and Goldstein, 1996) when using multifactor tests. Relying primarily on correlation and logistic regression analysis, the study found that five core factors drive the overall outcome of the multifactor test. The plaintiff must win the similarity factor to win the overall test, and in every opinion in which the plaintiff also wins the intent factor, it wins the overall test. The proximity of the goods and marketing channels, evidence of actual confusion, and the strength of the plaintiff's mark also play important roles in driving the outcome of the test. The study further showed that when judges find a likelihood of confusion, they tend to find that all the factors in the multifactor test support that result, while the factors do not tend to “stampede” in this way in opinions in which no likelihood of confusion is found. Relatedly, a finding that the defendant acted with a bad-faith intent to infringe tends to trigger stampeding of the factors all to support a likelihood of confusion. With respect to the *Abercrombie* spectrum of trademark distinctiveness, the study showed that judges make far less use of this spectrum than is generally thought, and that the inherent distinctiveness analysis appears to exert little influence over the court's overall assessment of the strength of the mark, which is guided instead by the court's assessment of the actual marketplace strength (or weakness) of the mark. Finally, the study made the controversial finding, discussed further below, that survey evidence is rarely presented in likelihood of confusion disputes and that when it is presented, judges rarely credit it.

Blum, Fox, Hayes, and Xu (2010) renewed the systematic content analysis of federal court opinions employing the multifactor test for the likelihood of consumer confusion by focusing on a 15-year period from 1994–2008 and on federal court opinions from one district, the Southern District of New York. Working from an original dataset of 206 opinions, they sought to determine whether the findings in Beebe's 2006 study would hold across time. Their

findings were largely consistent with Beebe's with respect to win rates, the relative importance of the various factors, and courts' treatment of the *Abercrombie* spectrum and inherent distinctiveness. However, though Beebe had found that Second Circuit district courts were somewhat less prone to stampeding the factors than other districts, the authors found no evidence at all that S.D.N.Y. courts stamped the factors when finding a likelihood of confusion or, more specifically, when finding that the defendant acted with a bad-faith intent to infringe. As for historical trends, the authors grouped the opinions into three-year baskets and reported a declining proportion over time of opinions that found a likelihood of confusion. The authors also considered whether an additional factor—whether the parties' goods were directly competing—could help predict the overall outcome of the test. They concluded that the competing goods factor (which they admitted is closely similar to the proximity of the goods factor) did no better than any other factors in predicting the outcome of the multifactor test.

In her study of Google search results discussed above, Ouellette (2014) analyzed the first page of search results to assess not only the distinctiveness of the asserted mark, but also the likelihood of confusion of the parties' marks. Addressing in detail many of the cases in her dataset and their Google search results, she noted the many cases in which the search results either supported a finding of a likelihood of confusion (e.g., because the search results for the defendant's mark contained hits referring to the plaintiff) or did not support a likelihood of confusion (e.g., because the search results for the defendant's mark contained no hits referring to the plaintiff). Overall, Ouellette identified 16 cases in which the Google search results and the court disagreed on the issue of likelihood of confusion. In only three of these, she argued, was the court correct rather than Google; and in each of these three cases, the defendant was making a permissible concurrent use of a mark in a limited geographical area—a special factual situation that Google search results are not equipped to account for. Ouellette suggested that Google search results, like traditional survey evidence, may serve as probative evidence in trademark litigation; but she was quick to emphasize that, like survey evidence, Google's search results must always be subject to careful assessment by the fact finder.

B. The Likelihood of Confusion and Sponsorship Confusion

Trademark law will generally find infringement when the defendant uses a trademark in such a way that consumers believe that even though the plaintiff did not produce the defendant's goods, the plaintiff has somehow sponsored, authorized, or endorsed them. So-called "sponsorship confusion" is controversial in trademark commentary. Lemley and McKenna (2010) in particular have advocated that trademark law adopt the principle that sponsorship confusion is actionable only when it is material to consumers' decision to purchase.

In light of the controversy surrounding sponsorship confusion, Kugler (2017) conducted an Internet consumer confusion survey to determine how often sponsorship confusion is material to consumers and in what ways it is material. He exposed consumers to various online advertisements patterned after Amazon listings for merchandising products for professional sports teams, well-known automobile brands, a well-known university, various city and governmental institutions (e.g., the New York Police Department), parody products, and a motion picture whose title was the subject of a claim of sponsorship confusion. Kugler adapted the standard *Ever-Ready* consumer confusion survey format (*Union Carbide v. Ever-Ready*, 531 F.2d 366 (1976)) to include additional questions testing for materiality. Specifically, he asked survey respondents: (1) who they thought was responsible for the quality of the product;

(2) whether they would have a better, worse, or unchanged opinion about the sponsor of the product if they had an adverse experience with the product; (3) assuming that the product was not produced by the entity to which the product referred, whether they would be more or less interested in purchasing the product; and (4) whether they would be willing to pay more for the product if it was or was not made with the permission or authorization of the entity to which the product referred. Kugler analyzed the responses of 1049 consumers, various subsets of whom were exposed to various subsets of the merchandising products at issue. Overall, he interpreted the data as showing that “across a wide range of products and domains, sponsorship confusion is rarely material to more than half of potential consumers.” On this basis, he concluded that adopting a materiality requirement would allow a variety of unauthorized third-party merchandising uses of well-known trademarks that the law currently prohibits. However, Kugler emphasized that the data show that in the context of professional sports brands in particular, sponsorship confusion is typically material to consumers, and so the adoption of a materiality requirement would not significantly affect merchandising practices in that context. Interestingly, he also asked respondents whether they thought it should be possible to make the advertised product without the permission of the entity referred to by the product. The data show that for most products tested, most respondents believed that permission should be obtained.

C. The Likelihood of Confusion in Context

In two large-scale empirical studies, David Franklyn and David Hyman (2013, 2014) focused on the likelihood of confusion in the context of keyword advertising. They designed and oversaw the administration of three surveys, two of which were conducted in 2010 and one in 2012, and each of which had about 1000 respondents. The surveys asked a wide array of questions designed to determine how consumers use search engines when searching for branded goods and how consumers perceive advertisements in the form of paid links that appear together with organic links on the results pages of leading search engines. The survey results indicate a surprisingly high degree of consumer uncertainty and confusion about search page architecture and the labelling of paid links, with a majority of respondents indicating that they pay no attention to where links are located on the search results page or whether the links appear in a shaded box (which is one way that search engines have tried to distinguish paid links). Franklyn and Hyman (2013) conclude that most consumers ignore labels identifying paid links, and thus that such labels fail to communicate whether content appearing on results pages is paid or unpaid. Nevertheless, the authors find in their survey results little evidence that consumer ignorance with respect to paid links actually results in forms of consumer confusion that are traditionally actionable under trademark law. While consumers may not understand why paid links appear or understand the difference between paid and unpaid links, only a small minority of consumers (15 percent, in the authors’ estimate) tend to incorrectly assume that paid links purchased by third parties are actually sponsored, endorsed, or somehow affiliated with the owners of the trademarks to which such paid links are keyed. Interestingly, however, the authors report that about half of those surveyed (excluding those who were unsure or without an opinion) believed that it was unfair or inappropriate for one firm to purchase the trademark of a competitor for use as a keyword triggering a paid link.

In a second study, which they conducted mainly in 2010, Hyman and Franklyn (2014) studied which entities were purchasing trademarks as keywords for purposes of keyword advertising. To do so, they systematically studied the search results generated when 2462

prominent trademarks were entered into three leading search engines. Among other data they collected, they classified the “first five paid and unpaid links” for each mark entered in each search engine according to whether the link referred to the trademark owner, a competitor, or various other entities. They found that firms purchased the trademarks of their competitors as keywords in only 6 percent of the paid links they studied. Instead, trademarks owners themselves account for a high proportion of keyword purchasing activity, apparently to ensure that their links appear at the top of the search results and perhaps to prevent their competitors from benefiting from paid links keyed to their trademarks. The authors also found a high degree of variation over time in the purchase of paid links and the domain names to which those links referred. Overall, they conclude that, notwithstanding legal attention paid to the phenomenon in the form of a spate of litigation and commentary in recent years, the risk of widespread abuse of keyword advertising is quite low.

DeRosia, Lee, and Christensen (2011) developed experimental evidence showing that in certain contexts, more sophisticated consumers may actually be more likely to experience confusion as to source than less sophisticated consumers. This runs counter to the intuition—and to the black-letter principle in trademark law—that the level of a consumer’s sophistication is inversely related to that consumer’s likelihood of confusion. The authors hypothesized, in essence, that if a competitor of a senior mark extends its brand into a new product category, then consumers of that senior mark are more likely to be confused by a junior mark in the new product category that is identical or similar to the senior mark. The authors further hypothesized that more sophisticated consumers are more likely to be confused. Due to their heightened sophistication, such consumers more readily update their mental schemas upon exposure to the competitor’s brand extension and are more prone to assume, upon exposure to the junior mark, that the senior mark has also engaged in the extension of its brand. In order to explore these hypotheses, the authors conducted an experiment which exposed participants to a series of advertisements, some of which announced the extension of certain automobile brands into laptop computers, and tested the participants for their “enduring involvement” in the product category of automobiles and their “need for cognition” —that is, “an individual’s enduring tendency to engage in and enjoy effortful cognitive endeavors.” The experimental results confirmed the authors’ hypotheses.

D. The Role of Survey Evidence

Trademark law scholars have devoted a great deal of attention to the role of survey evidence in the likelihood of confusion analysis, and for good reason. Ford (2012; 2013) tracked the use of survey evidence in reported cases of federal trademark litigation from 1946–2012 and found a significant increase in the use of such evidence in connection with some question of fact in trademark disputes, with it appearing very rarely in the reported case law in the mid-twentieth century (e.g., in only about six reported opinions per year for the period 1961–75) and now appearing relatively more frequently (in about 45 reported opinions per year over the period 2006–12). However, Beebe’s 2006 empirical study of the multifactor test suggested that despite conventional wisdom to the contrary (and the insistent claims of survey experts), survey evidence of the likelihood of consumer confusion is not commonly presented in trademark litigation; and more importantly, that when it is presented, courts rarely give it any weight, particularly if the evidence runs contrary to the court’s ultimate finding on the issue of the likelihood of confusion. Of the 331 opinions Beebe studied, only 20 percent discussed survey evidence and only 10 percent credited that evidence (Beebe, 2006).

Other quantitative studies have found that survey evidence is more frequently presented than Beebe's 2006 study suggested. As part of a larger study of the consumer in trademark law, Austin (2004) evaluated ten years of trademark infringement case law from May 1993 through May 2003 to find that survey evidence "is before the court around 57.4 percent of the time," but that courts gave weight to that evidence in only 35.2 percent of the cases studied. Austin characterizes his findings as showing a "low use of survey data in trademark cases."

Also contrary to Beebe's findings, Sarel and Marmorstein (2009) studied trademark infringement cases from 2001–06 and found that survey evidence was considered in 34.1 percent of these cases. To develop their dataset, the authors included only opinions (1) "in which the infringement plaintiff possessed an undisputed, valid trademark," (2) which resulted in a dispositive judgment on the likelihood of confusion, and (3) which conveyed "sufficient information about the basic issues of the case and evidence presented." Furthermore, the authors excluded cases that "were dismissed on technicalities or that focused primarily on other legal issues" outside of the likelihood of consumer confusion. The dataset thus consisted of 126 cases. The authors then coded the cases for a variety of characteristics, including whether, in the authors' view, the parties' marks were "more similar" or "less similar." The authors found overall no statistically significant difference in plaintiff win rates between cases in which the plaintiffs presented "actual confusion or survey evidence" and those in which plaintiffs presented no such evidence. However, and quite interestingly, the authors do show that in the cases in which the parties' marks were coded as "less similar," plaintiffs enjoyed a substantially higher win rate (85.7 percent) when they presented survey evidence on the issue of the likelihood of consumer confusion and that evidence was admitted than when they did not present survey evidence (27.3 percent) or they did and the evidence was not admitted (0 percent). (The authors do not specify how many cases in their dataset were coded as less similar, so the strength of this finding is not clear).

Bird and Steckel (2012) expanded on Beebe's initial dataset of 311 opinions from 2000–04 to study the treatment of survey evidence in 533 reported federal district court opinions from 2000–06 that made substantial use of a multifactor test for the likelihood of consumer confusion. After coding the additional two years of opinions according to Beebe's protocol, the authors found that only 16.6 percent of the opinions they studied addressed survey evidence and thus that "survey evidence is used infrequently" in trademark litigation. However, in opinions in which survey evidence was considered, it did sometimes appear to influence the outcome of the confusion analysis. The authors used regression analysis to conclude that "[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 confusing finding." Furthermore, the authors suggested that surveys may play significant roles that would not be apparent from published opinions, specifically in providing guidance to plaintiffs considering filing suit or in giving plaintiffs leverage in settlement negotiations.

Diamond and Franklyn (2014) significantly advanced the empirical study of survey evidence in trademark litigation by looking outside of reported published opinions to ask trademark litigators what their experience had been with survey evidence. In 2013, the authors sent invitations to participate in a survey about "consumer perception surveys" to all members of the International Trademark Association. Of the 465 respondents, 335 described themselves as practicing attorneys (of whom two were excluded from the tabulation of results because they also described themselves as survey consultants). More than half of the remaining 333 practicing attorneys indicated that they had commissioned at least one survey,

and of the 145 attorneys from the U.S. who indicated that they had practiced for at least eight years, 96 (or 61 percent) reported that they had commissioned at least one survey. Indeed, for these 96 attorneys, the average number of surveys each had commissioned so far in his or her career was 11.8 surveys. Diamond and Franklyn present further statistics strongly suggesting that only a small proportion of these surveys were actually presented at trial and that in the experience of the attorneys surveyed, consumer perception surveys can play a very significant role in the lead-up to trial and in settlement negotiations.

Outside of the U.S., Huang, Weatherall, and Webster (2012) conducted an empirical study of the use of survey evidence in Australian federal and state trademark and passing off litigation. They collected 353 trademark and passing off cases from the LexisNexis database for the period from 1990 through April 1, 2010. Finding only 33 cases that discussed survey evidence, they identified an overall downward trend in the use of survey evidence over that time. In a careful probit regression analysis incorporating a variety of case-specific factors, the authors concluded that they could not reject the null hypothesis that survey evidence had no effect on case outcomes for the cases studied. However, they muster a variety of findings to show that surveys had at best a “limited impact” on outcomes.

V. THE LIKELIHOOD OF DILUTION TEST

Dilution describes the process by which certain kinds of non-confusing uses of a trademark by third parties can diminish that mark’s distinctiveness of source or distinctiveness from other marks. For example, if a small corner bodega calls itself “Apple Market,” it is not at all likely that consumers will believe the bodega is somehow affiliated with the maker of iPhones. Thus, the bodega’s use does not create a likelihood of consumer confusion as to source and no cause of action will succeed on that ground. However, if many shops of many different kinds begin to use the term “Apple” in their names, though consumers may continue to realize that there is no connection between such shops and the high-technology firm, the distinctiveness of the high-technology firm’s mark may be impaired. Upon hearing “Apple” in a commercial context, consumers may have to think for a moment to determine to which firm the term is referring, and this increases consumer search costs (or so goes the orthodox story in trademark law). Furthermore, as more and more firms use the word “Apple,” the distinctiveness of the mark from other marks diminishes. It becomes a commonplace commercial designation for a wide variety of firms. Its advertising power therefore declines.

A. Anti-dilution Law in the Courts

Anti-dilution protection has remained a highly controversial area in trademark law. In an ambitious study, Long (2006) was the first empirically to confirm the conventional wisdom that courts are hostile to the anti-dilution cause of action. Long developed a dataset of 344 cases that considered a federal anti-dilution cause of action drawn from all reported federal opinions on Westlaw from the January 16, 1996 effective date of the Federal Trademark Dilution Act of 1995 (FTDA), which established anti-dilution protection at the federal level, through July 16, 2005. Her data show a substantial and steady decline over time in the success rate of dilution claims, with approximately 54 percent of such claims granted in 1996 but only 12 percent granted in the first half of 2005. This decline is greater than the decline in win rates over the same period for non-dilution trademark claims. The author also used data from the

Administrative Office of the United States Courts and the Public Access to Court Electronic Records (PACER) database to study cases not reported in Westlaw. She focused on the ten federal districts with the greatest number of trademark complaints and studied, for the period January 16, 1996 through 1999, every dilution-related complaint that was available. For the period 2000 through July 16, 2005, she took a random sample. This resulted in a dataset of 732 dilution filings. Long found similar declines in dilution win rates in the unreported case law. She theorizes overall that “[a]lthough judges are not justifying their actions in efficiency terms, over time they are generally denying enforcement in cases where enjoining unauthorized use of the trademark would reduce social welfare.” Though dilution claims appear to have been largely unsuccessful in the federal courts over the period she studied, Long nevertheless recognizes that dilution could be a “powerful bargaining chip” in cease-and-desist stage or settlement negotiations.

In a much smaller-scale study, Beebe (2008) conducted a systematic content analysis of all reported federal court opinions that analyzed an issue in anti-dilution law during the one-year period following the October 6, 2006 effective date of the new Trademark Dilution Revision Act (TDRA), which replaced the FTDA. Using Westlaw and Lexis, Beebe developed a dataset of 85 opinions. Perhaps unsurprisingly, the data showed that many courts were unaware of the new law and thus improperly applied the old FTDA, or that even if they were aware of the new law, they continued to apply precedent based on the FTDA. Some courts apparently unknowingly applied bits and pieces of both the old law and the new. The main finding of the study was that the anti-dilution cause of action is largely redundant of the infringement cause of action. Of the 64 opinions studied that analyzed both a confusion and a dilution cause of action, 89 percent reached the same outcome under both. No opinion found no confusion but then went on to find dilution. In only one opinion did the dilution analysis appear to drive the outcome of the case. In ongoing work, Beebe (2017) updates these results for the ten-year period following the effective date of the TDRA. The data show that the dilution cause of action remains essentially redundant of the confusion cause of action and plays no significant role in driving the outcomes of federal trademark infringement case law.

As mentioned above, Sheff (2014) undertook a comprehensive study of the effect—or more precisely, the lack of effect—of federal anti-dilution law on registration practice at the USPTO. Sheff conducted three different empirical analyses. First, he developed and hand-coded a dataset of all 453 Trademark Trial and Appeal Board (TTAB) dispositions of dilution claims from the January 16, 1996 effective date of the FTDA through June 30, 2014. He found only three TTAB cases over that 18-year period in which anti-dilution claims made any difference to the outcome of a TTAB adjudication; and in one of these, Sheff argues, a likelihood of confusion claim could have been used to reach the same outcome, but having found dilution, the TTAB declined to consider the confusion claim. The other two cases were free speech cases with highly controversial findings of dilution. Second, Sheff used the TCFD discussed above to study trends in the application or registration rates of identical marks by different parties in different product categories. He found no significant change in such rates that would suggest that anyone had changed their behavior in response to the advent of federal anti-dilution law. Third, Sheff studied third-party applications to register famous marks. To do so, he developed a list of 86 clearly famous brands and studied all third-party applications for such trademarks (or close approximations of such trademarks) filed on or after January 1, 1981. He found that such applications are “extremely uncommon” and that, in any case, there was no change in the annual rate of filing of such applications that might suggest some effect of the advent of federal anti-dilution law. The author’s overall conclusion is compelling:

“dilution appears to have been largely (though perhaps not entirely) a failure at the PTO.” The primary effect of anti-dilution rights “appears to be imposing increased costs on those who have business with the trademark system, with little if any apparent benefit to justify those costs.”

B. Dilution Over Time

A long-standing challenge in anti-dilution law has been to show an example of a single famous trademark that has actually suffered from dilution. Many famous marks have eventually collapsed into generic terms, but it is far more difficult to identify a famous mark that has maintained its distinctiveness of source, but nevertheless seen that distinctiveness significantly diluted over time by third-party non-confusing uses of the mark.

Two important articles by Brauneis and Heald (2011a; 2011b) have taken up the considerable challenge of studying trademark dilution over time. In the first (2011a), the authors developed a list of 131 historically famous brand names based in part on the list published in the early twentieth century by Hotchkiss and Franken (1923) and updated in the late century by Golder (1997). They then examined business names used in the white pages telephone directories of Chicago, Philadelphia, and Manhattan in the years 1940, 1960, 1980, 1990, 2000, and 2010. Among the authors’ most interesting general findings is that from 1960–2010, unauthorized third-party uses of famous marks declined by 54 percent across the three cities. The authors closely considered many economic, cultural, and legal factors that may have contributed to this decline. They suggested that the introduction of anti-dilution protection and increased anti-confusion protection partially explains the decline.

In the second article, Brauneis and Heald (2011b) moved beyond white pages telephone directories to study the appearance of third-party uses of famous brands names in a wide variety of sources. Specifically, they examined: (1) state corporate name and LLC databases over various periods of time from six states (three that established anti-dilution protection long before the 1996 advent of federal anti-dilution protection and three that offered no state-level anti-dilution protection before 1996); (2) advertisements appearing from 1925–2010 in the *New York Times*, *Washington Post*, and *Wall Street Journal*; (3) all reported dilution litigation in the federal courts from 1946 apparently through 2010; and (4) Westlaw’s state trademark registration database and the Principal Register of federal trademark registrations at the USPTO. At the same time that they greatly expanded their corpus of sources where unauthorized third-party uses might appear, they narrowed the number of historically famous brand names on which they focused from 131 to 33. The authors presented a host of interesting findings with respect to certain brand names in certain states. Among their main overall findings was that third parties commonly make unauthorized uses of famous marks as trade names, but rarely do so for products. The authors further showed significant declines over time in such unauthorized third-party uses; and as they did in their first article (2011a), the authors suggested that the emergence of anti-dilution protection at the state and federal levels likely contributed to these declines.

C. Experimental and Survey Evidence of Dilution

Another considerable challenge in anti-dilution law has been to design an experiment or survey method that would show dilution. Morrin and Jacoby (2000) developed a series of experiments that remains the leading effort in this regard. They began by hypothesizing that consumers exposed to diluting stimuli are more likely to commit brand recognition errors and

to exhibit slower brand recognition times than consumers unexposed to such stimuli. To test this hypothesis, the authors exposed 64 study participants to six print advertisements in random order. For some of the participants, some of these advertisements conveyed diluting information—that is, the advertisements associated a famous brand with a product or service with which the brand was not traditionally associated. The study participants then completed a computer task in which they were exposed to a brand name (or a product category/attribute) on the computer display followed by a product category/attribute (or a brand name). The participants were instructed to indicate as quickly as possible by pressing a key for yes or no whether the two words “represented a match.” The authors found that response accuracy with respect to certain brand names was significantly lower for participants exposed to stimuli diluting of those brands. The authors further found that response times were significantly higher with respect to certain brands for participants exposed to stimuli diluting of those brands. However, the authors found no significant change in response times for the particular brand (Hyatt for hotel services) with which the study participants were most familiar.

Morrin and Jacoby (2000) then conducted a second study designed to test participants’ product category recall for various brands after exposure to diluting stimuli, some of which exposed certain participants to diluting uses in related product categories, while others exposed other participants to diluting uses in unrelated product categories. The authors found that participants’ category recall with respect to certain brands declined after exposure to stimuli diluting of those brands, with more pronounced declines in recall following exposure to diluting uses from unrelated product categories than from related product categories. Importantly, however, Morrin and Jacoby found that brands of especially high fame showed no declines in category recall. They conclude: “It appears that very strong brands are immune to dilution because their memory connections are so strong that it is difficult for consumers to alter them or create new ones with the same brand.”

Morrin, Lee, and Allenby (2006) developed a hierarchical Bayes associative network model for consumers’ memory of a brand and tested how exposure to diluting stimuli may impact that associative network. They also sought to determine how contextual factors, and in particular, the presence of consumer confusion, may affect brand recall. The authors exposed 212 participants to a booklet containing various real-world brand logos below which appeared their product categories. Some participants were given booklets that contained real-world logos of companies that used very similar trademarks (e.g., Bass for shoes and Bass for ale). After a distracter task, the participants were provided with a list of certain brands of interest and given one minute to record which product categories came to mind for each brand name. The participants were also asked to assess on a nine-point scale various contextual factors, such as their level of familiarity with the various brands included in the booklet and their level of confusion as to source with respect to the various brands. The authors analyzed, among other things, brand exclusive recall (i.e., the proportion of consumers who, when presented with one of the brand names used by two different companies, recalled only the product category of the company that first used the brand name). The authors conclude that “a single exposure to diluting brand stimuli is found to have a damaging effect on brands, reducing brand exclusive recall by about a third, on average.” Once again, however, the authors show that highly famous brand names were not significantly impacted by the diluting stimuli used in the experiment. They explain:

The legal implication of this result is that if the level of fame required by the courts for application of the FTDA is held extremely high, then few of the brands to which it applies may ever be able to demonstrate harm, at least from single instances of trademark dilution.

The study is also of great interest because its results suggest that consumers' level of confusion with respect to two very similar brands may impact the degree to which the junior user's use dilutes the senior user's brand.

In a wide-ranging and remarkably ambitious collection of five experiments, Pullig, Simmons, and Netemeyer (2006) used three different measurement methodologies—response latency, aided recall, and simulated choice—to study how junior brands affect identical or very similar senior brands when the brands are in related or unrelated product categories and are perceived to possess similar or dissimilar attributes. The results of the experiments suggested that junior brands will not dilute identical or very similar senior brands when the brands are in the same product category and possess similar attributes. However, if a junior brand possesses dissimilar attributes and is used in a product category similar to that of an identical or very similar senior brand, the junior brand's use will dilute the attribute associations of the senior brand. Furthermore, if a junior brand possesses dissimilar attributes and is used in a dissimilar product category, its use will dilute both the category and attribute associations of the senior brand. Finally, with respect to simulated choice, the authors showed that when a junior brand is dissimilar to a senior brand along the dimensions of brand attributes and product category, this can suppress in certain conditions consumers' consideration and choice of the senior brand.

Notwithstanding many experimental attempts to show trademark dilution, many trademark scholars remain unconvinced that these experiments have succeeded in demonstrating that dilution represents a real problem in the marketplace. Tushnet (2008) presented the most prominent of these critiques. She questioned, among other things, whether the slight decreases in response times recorded in the experiments actually translate into any significant harm to famous brands. She further questioned whether lab experiments can adequately recreate the full context of the purchasing experience and decision, particularly in light of the fact that the consumption context may altogether dispel any significant impairment of a famous mark's attribute or product category associations conventionally labelled "dilution."

Beebe, Germano, Sprigman, and Steckel (2017) conducted a series of experiments to test whether the protocols used by Morrin and Jacoby (2000) and Pullig, Simmons, and Netemeyer (2006), which involved response latency, properly measured dilution. Beebe, Germano, Sprigman, and Steckel found, in essence, that respondents exposed to any diluting stimuli slowed their response times to questions involving nearly any brand, even if the brand was not the target of the diluting stimuli. They proposed that unfamiliar diluting stimuli, such as an advertisement for Mercedes toothpaste, cause respondents to experience surprise and to adopt a stance of guardedness toward subsequent test questions, thus increasing response times. On this basis, they contended that previous response time studies used the wrong control, in that they exposed a treatment group to diluting stimuli and the control group to no diluting stimuli. The authors showed that when the control group is exposed to diluting stimuli that targeted brands other than those that were the focus of the treatment, there is no significant difference in response times between the treatment and control groups. They concluded that current response time methods of testing for dilution are unable to show any evidence of dilution.

VI. CONCLUSION

In just two decades, the empirical study of trademark law has rapidly developed into a sophisticated and wide-ranging area of inquiry. The USPTO's recent release and ongoing updating of the TCFD represents an especially significant achievement that will undoubtedly continue to inspire scholarly work in the area. It is hoped that other leading intellectual property offices, particularly the EUIPO, will also publicly release their data for scholarly examination. Further empirical work also remains to be done in a number of areas relating to trademark litigation, including how courts adjudicate defenses to infringement and what remedies courts provide for infringement.

REFERENCES

- Austin, Graeme W. 2004. "Trademarks and the Burdened Imagination", 69 *Brooklyn Law Review* 827–922.
- Beebe, Barton. 2006. "An Empirical Study of the Multifactor Tests for Trademark Infringement", 95 *California Law Review* 1581–654.
- Beebe, Barton. 2008. "The Continuing Debacle of U.S. Antidilution Law: Evidence from the First Year of Trademark Dilution Revision Act Case Law", 24 *Santa Clara Computer & High Technology Law Review* 449–67.
- Beebe, Barton. 2012. "Is the Trademark Office a Rubber Stamp?", 48 *Houston Law Review* 751–77.
- Beebe, Barton. 2018. "The Continuing Redundancy of U.S. Trademark Antidilution Law: Evidence from Ten Years of Trademark Dilution Revision Act Case Law", Working Paper.
- Beebe, Barton, and Jeanne Fromer. 2018. "Are We Running Out of Trademarks?: An Empirical Study of Trademark Depletion and Congestion", forthcoming *Harvard Law Review*.
- Beebe, Barton, and Roy Germano, Christopher Jon Sprigman, and Joel Steckel, "Testing for Trademark Dilution in Court and the Lab", Working Paper.
- Bird, Robert C., and Joel H. Steckel. 2012. "The Role of Consumer Surveys in Trademark Infringement: Empirical Evidence from the Federal Courts", 14 *University of Pennsylvania Journal of Business Law* 1013–53.
- Blum, Kevin, Ariel Fox, Christina J. Hayes, and James (Hanjun) Xu. 2010. "Consistency of Confusion? A Fifteen-Year Revisiting of Barton Beebe's Empirical Analysis of Multifactor Tests for Trademark Infringement", *Stanford Technology Law Review* 3–45.
- Carpenter, Megan M., and Mary Garner. 2015. "NSFW: An Empirical Study of Scandalous Trademarks", 33 *Cardozo Arts & Entertainment Law Journal* 321–65.
- DeRosia, Eric D., Thomas R. Lee, and Glenn L. Christensen. 2011. "Sophisticated but Confused: The Impact of Brand Extension and Motivation on Source Confusion", 28 *Psychology & Marketing* 457–78.
- Diamond, Shari Seidman, and David J. Franklyn. 2014. "Trademark Surveys: An Undulating Path", 92 *Texas Law Review* 2029–73.
- Ford, Gerald L. 2012. "Survey Percentages in Lanham Act Matters", in Shari Seidman Diamond and Jerre B. Swan, eds., *Trademark and Deceptive Advertising Surveys: Law, Science, and Design*. American Bar Association.
- Ford, Gerald L. 2013. "Survey Evidence", *McCarthy Law Symposium*, February 29, 2013.
- Franklyn, David J., and David A. Hyman. 2013. "Trademarks as Search Engine Keywords: Much Ado About Something?", 26 *Harvard Journal of Law & Technology* 481–543.
- Gerhardt, Rebecca, and Jon P. McClanahan. 2013. "Do Trademark Lawyers Matter?", 16 *Stanford Technology Law Review* 101–41.
- Gigerenzer, Gerd, and Daniel G. Goldstein. 1996. "Reasoning the Fast and Frugal Way: Models of Bounded Rationality", 103 *Psychological Review* 650–69.
- Golder, Peter N. 2000. "Historical Method in Marketing Research with New Evidence on Long-Term Market Share Stability", 37 *Journal of Marketing Research* 156–72.
- Graham, Stuart, Galen Hancock, Alan Marco, and Amanda Fila Myers. 2013. "The USPTO Trademark Case Files Dataset: Descriptions Lessons, and Insights", U.S. Patent and Trademark Office.
- Helmets, Christian, Mark Rogers, and Philipp Schautschick. 2011. "Intellectual Property at the Firm-Level in the UK: The Oxford Firm-Level Intellectual Property Database", Oxford University Department of Economics Discussion Paper No. 546.
- Hotchkiss, George Burton, and Richard B. Franken. 1923. *The Leadership of Advertised Brands: A Study of 100 Representative Commodities Showing the Names and Brands that are Most Familiar to the Public*. Garden City, NY: Doubleday, Page & Company.

- Huang, Vicki, Kimberlee Weatherall, and Elizabeth Webster. 2012. "The Use of Survey Evidence in Australian Trade Mark and Passing Off Cases", in Andrew T. Kenyon, Megan Richardson, and Wee Loon Ng-Loy, eds., *The Law of Reputation and Brands in the Asia Pacific*. Cambridge University Press.
- Hyman, David A., and David J. Franklyn. 2014. "Trademarks as Search-Engine Keywords: Who, What, When?", 92 *Texas Law Review* 2117–50.
- Kugler, Matthew. 2017. "The Materiality of Sponsorship Confusion", 50 *U.C. Davis Law Review* 1911.
- Kur, Annette, and Roland Knaak. 2011. "Study on the Overall Functioning of the European Trade Mark System", Max Planck Institute for Intellectual Property and Competition Law.
- Lee, Thomas R., Eric D. DeRosia, and Glenn L. Christensen. 2009. "An Empirical and Consumer Psychology Analysis of Trademark Distinctiveness", 41 *Arizona State Law Journal* 1033–109.
- Lemley, Mark A., and Mark McKenna. 2010. "Irrelevant Confusion", 62 *Stanford Law Review* 413–54.
- Long, Clarissa. 2006. "Dilution", 106 *Columbia Law Review* 1029–78.
- Morrin, Maureen, and Jacob Jacoby. 2000. "Trademark Dilution: Empirical Measures for an Elusive Concept", 19 *Journal of Public Policy & Marketing* 265–76.
- Morrin, Maureen, Jonathan Lee, and Greg M. Allenby. 2006. "Determinants of Trademark Dilution", 33 *Journal of Consumer Research* 248–57.
- Ouellette, Lisa Larrimore. 2014. "The Google Shortcut to Trademark Law", 102 *California Law Review* 351–407.
- Pullig, Chris, Carolyn J. Simmons, and Richard G. Netemeyer. 2006. "Brand Dilution: When Do New Brands Hurt Existing Brands?", 70 *Journal of Marketing* 52–66.
- Sarel, Dan, and Howard Marmorstein. 2009. "The Effect of Consumer Surveys and Actual Confusion Evidence in Trademark Litigation: An Empirical Assessment", 99 *Trademark Reporter* 1416–36.
- Schautschick, Philipp, and Christine Greenhalgh. 2015. "Empirical Studies of Trade Marks—The Existing Economic Literature", 25 *Economics of Innovation and New Technology* 358–90.
- Sheff, Jeremy N. 2014. "Dilution at the Patent and Trademark Office", 21 *Michigan Telecommunications & Technology Law Review* 79–140.
- Simon, Dan. 2004. "A Third View of the Black Box: Cognitive Coherence in Legal Decision Making", 71 *University of Chicago Law Review* 511–86.
- Tushnet, Rebecca. 2008. "Gone in Sixty Milliseconds: Trademark Law and Cognitive Science", 86 *Texas Law Review* 507–68.
- United States Patent and Trademark Office. 2015. "Trademark Case Files Dataset", available at www.uspto.gov/learning-and-resources/electronic-data-products/trademark-case-files-dataset-0.
- von Graevenitz, Georg. 2013. "Trade Mark Cluttering: Evidence from EU Enlargement", 65 *Oxford Economic Papers* 721–45.
- von Graevenitz, Georg, Christine Greenhalgh, Christian Helmers, and Philipp Schautschick. 2012. "Trademark Cluttering: An Exploratory Report", U.K. Intellectual Property Office.

Cases

- Abercrombie & Fitch Co. v. Hunting World, Inc.* (1976), 537 F.2d 4.
- Triangle Publ'ns, Inc. v. Rohrlich* (1948), 167 F.2d 969.
- Union Carbide v. Ever-Ready* (1976), 531 F.2d 366.