1-1-1999

The Predictability of Punitive Damages Awards in Published Opinions, the Impact of BMW v. Gore on Punitive Damages Awards, and Forecasting Which Punitive Awards Will Be Reduced

Theodore Eisenberg
Cornell Law School, ted-eisenberg@lawschool.cornell.edu

Martin T. Wells
Cornell University, mtw1@cornell.edu

The Predictability of Punitive Damages Awards in Published Opinions, the Impact of *BMW v. Gore* on Punitive Damages Awards, and Forecasting Which Punitive Awards Will Be Reduced

Theodore Eisenberg & Martin T. Wells†

This article assesses the relation between compensatory damages and punitive damages in cases leading to published opinions and BMW v. Gore's *impact on the patterns of punitive damages awards in these opinions. We find that punitive damages awards are considerably higher in cases leading to published opinions than in trial level cases. But the correlation between compensatory and punitive awards found in trial level data persists in published opinions and is all but indistinguishable from the correlation in trial level data. We find no significant difference in the pattern of awards before and after* BMW *and no significant difference in the rate at which courts order a reduction in punitive damages awards. We also find that the mass of trial level awards provides a powerful tool for predicting the outcome of judicial review of punitive damages awards.*

TABLE OF CONTENTS

† Henry Allen Mark Professor of Law, Cornell Law School; Associate Professor, Department of Social Statistics, Cornell University. The authors thank an anonymous referee for comments.

In *BMW of North America, Inc. v. Gore*,[1] the Supreme Court relied in part on "a breathtaking 500 to 1"[2] punitive to compensatory damages ratio to invalidate, on federal constitutional grounds, an Alabama punitive damages award. Several courts have relied on *BMW* to reduce punitive damages awards.[3] And pundits, including us, have taken to describing *BMW*'s effect on ratios of punitive to compensatory damages that will survive judicial scrutiny.[4] This article assesses the relation between compensatory damages and punitive damages in cases leading to published opinions and *BMW*'s impact on the patterns of punitive damages awards in these opinions.

The relation between compensatory and punitive awards has been studied in trial-court-level data sets,[5] but it has not been analyzed in the mass of published opinions. One interesting question is whether the striking correlation between compensatory and punitive

[1] 116 S Ct 1589 (1996).

[2] 116 S Ct at 1603.

[3] See Evan M. Trager, BMW v. Gore, *One Year Later: The Road to Nowhere or Meaningful Guidance?*, 1997 Andrews Hazardous Waste Litig. Rep. 32711 (1997 ANHWLR 32711 on Westlaw) ("Trager, *One Year Later*").

[4] See Theodore Eisenberg & Martin T. Wells, *Punitive Awards After* BMW, *a New Capping System, and the Reported Opinion Bias*, 1998 Wisc L Rev 387 ("Eisenburg & Wells, *After BMW*"); Trager, *One Year Later* (cited in note 3); Sabrina C. Turner, Note, *The Shadow of* BMW of North America, Inc. v Gore, 1998 Wis L Rev 427, 450-56 ("Note, *Shadow of BMW*").

[5] See Theodore Eisenberg, John Goerdt, Brian Ostrom, David Rottman, and Martin T. Wells, *The Predictability of Punitive Damages*, 26 J Leg Studies 623 (1997) ("Eisenberg, et al, *Predictability of Punitive Damages*"); Stephen Daniels & Joanne Martin, *Civil Juries and the Politics of Reform* (Northwestern Univ Press, 1995); Erik Moller, *Trends in Civil Jury Verdicts Since 1985* (RAND, 1996).

awards in trial-level data[6] extends to the filtered set of cases that leads to published opinions. We use a substantial sample of pre- and post-*BMW* published opinions to explore the relation between punitive and compensatory awards.

After describing the relation between punitive and compensatory awards in published opinions, we consider two topics stemming from that relation. First, we study *BMW*'s effect on courts' treatment of punitive awards. Straightforward hypotheses in light of *BMW* are that punitive awards should be lower than before *BMW* and that the ratio of punitive awards to compensatory awards should have decreased. Second, we explore the pattern of review of awards. Since *BMW* was the first Supreme Court case to strike a punitive damages award on federal constitutional grounds, a further plausible hypothesis is that, after *BMW*, courts should be more willing to strike punitive damages awards as excessive than they were before *BMW*.

We find that punitive damages awards are considerably higher in cases leading to published opinions than in trial level cases, which usually do not lead to published opinions. But the correlation between compensatory and punitive awards found in trial level data persists in published opinions. Indeed, the correlation between compensatory and punitive awards in published opinions is all but indistinguishable from their correlation in trial level data.

We find no significant difference in the pattern of awards before and after *BMW* and no significant difference in the rate at which courts order a reduction in punitive damages awards. We do find meaningful differences in the ratio of punitive to compensatory awards before and after *BMW*. But the change occurs at the low end of the ratio distribution. Although one would have expected *BMW* to reduce mostly high ratio awards, these awards increased after *BMW*. Despite claims of *BMW*'s impact, it may be that not enough time has elapsed for *BMW* to exercise significant influence on the pattern of punitive awards.

We also find that statistical analysis of the mass of trial level awards provides a powerful tool for predicting the outcome of judicial review of punitive damages awards. Punitive-compensatory damages ratios that are beyond the upper 95 percent prediction band, as determined by trial level data, are much more likely to be struck in published opinions than are awards with ratios that are within the trial level upper 95 percent prediction band. The predictive power of a model based on the trial level data successfully classifies the outcome of over 90 percent of published opinion cases and exceeds

[6] Eisenberg et al, *Predictability of Punitive Damages* (cited in note 5).

the predictive power of models based on arbitrarily chosen multipliers, such as punitive awards being limited to three times compensatory awards. The predictive power of a system based on the greater of $500,000 or ten times the compensatory awards is about the same as one based on the trial level data.

Part I of this article describes our methodology and reports descriptive statistics. Part II examines the pattern of punitive and compensatory awards in published opinions from mid-1995 to mid-1997. Part III explores whether *BMW*, which was decided in May 1996, has had a detectable influence on the pattern of awards. Part IV illustrates how trial level data can be used to forecast which punitive awards will be reduced.

I. BACKGROUND, CASE SEARCH METHODOLOGY, AND DESCRIPTIVE STATISTICS

A. What Should the Relation Between Punitive and Compensatory Awards Look Like?

We sketch elsewhere the expected pattern of punitive awards.[7] Since this article focuses primarily on the relation between compensatory and punitive awards, we do not review theories of the frequency of punitive awards. Instead we briefly review the expected relation between the two kinds of awards.

A punitive award's level ought to, and often does, relate to the harm the defendant caused. The level of harm caused determines to some extent society's view of the egregiousness of the defendant's misbehavior. Because the compensatory award measures the degree of harm, it should correlate with the level of punitive damages, all other things being equal.

Other influences on punitive awards may include the defendant's financial status, the subject area of a case, the case's locale, and the defendant's corporate status.[8] But prior studies suggest the level of harm, as measured by compensatory damages, dwarfs other influences on the level of punitive awards.[9] We therefore do not emphasize these other possible influences on punitive award levels in this study. We are primarily interested in whether the relation between compensatory and punitive awards discovered in trial level data emerges in appellate cases and whether *BMW* affected that relation.

[7] See id

[8] Id at 626-632.

[9] Id at 628.

B. Search Methodology Used to Find Cases

This study uses a Westlaw search to identify cases with useable information about compensatory and punitive awards. The search is:

"sy((punitive exemplary) w/3 damages) & [date restrictions]."

The "sy" parenthetical limits the search to the text of the synopsis that Westlaw provides for each case. We searched Westlaw's Allfeds and Allstates databases for one year prior to *BMW*, extending back to May 20, 1995, and through mid-July 1997, somewhat more than one year after *BMW*. The search does not cover the text of the entire opinion. Thus, cases that discuss punitive damages, but for which Westlaw does not include the words "punitive damages" or "exemplary damages" in the synopsis, are not included in the study.

It is important to distinguish the case sample generated by this study's search from the sample used in two prior punitive damages articles. The Westlaw search used in this study differs from that used in our previous study of punitive damages awards opinions. The previous study required, in effect, that a case assessing a punitive damages award include a citation to *BMW v. Gore*.[10] The previous study thus included cases likely to discuss reducing a punitive award on the basis of *BMW*. But it did not include many cases in which no constitutional challenge was made to the punitive award, which are much less likely to have cited the *BMW* opinion. And it did not include cases that ordered awards reduced without citing *BMW*.[11] We refer to the previous *BMW*-based search as the "*BMW* search." One cannot detect a useful set of punitive damages cases decided prior to *BMW* using the *BMW* search. One needs a broader search such as the synopsis-based search used here.[12]

The instant study yields a case sample that also differs from that in our prior study of the predictability of punitive damages in trial level cases.[13] That sample, described elsewhere,[14] consisted solely of

[10] Eisenberg & Wells, *After BMW* at 408 (cited in note 4).

[11] *Proctor v Davis*, 682 NE2d 1203 (Ill Appl Dist), appeal denied, 682 NE2d 1146 (Ill 1997); *Maiorino v Schering-Plough Corp.*, 695 A2d 353 (NJ Super AD), certificate denied, 704 A2d 19 (NJ 1997).

[12] A possible broader search would span the text of opinions for the words punitive or exemplary. Such a search returns enormous numbers of cases that do not report useable punitive and compensatory awards.

[13] Eisenberg et al, *Predictability of Punitive Damages* (cited in note 5).

[14] Id at 632. For a complete description of the data, see *Inter-University Consortium for Political and Social Research, Civil Justice Survey of State Courts, 1992: [United States]*, ICPSR 6587 (First ICPSR ed., April 1996).

trial level cases tried to juries. Inclusion in that sample did not depend on publication of an opinion. Briefly, the dataset consisted of one year of punitive damages jury trial awards in 45 urban trial courts concluded in 1991-1992. The data are part of the Civil Trial Court Network ("CTCN"), a project of the National Center for State Courts and the Bureau of Justice Statistics. The CTCN obtains its data directly from court clerks' offices. It covers state courts of general jurisdiction in a random sample consisting of 45 of the 75 most populous counties in the United States. The 75 counties include approximately 33 percent of the 1990 U.S. population.[15] The CTCN jury trial data cover fiscal 1991-1992 (July 1 to June 30). We refer to the CTCN trial level data as the "trial level sample."

Our new Westlaw search yielded several hundred cases for possible inclusion in the study. Many of the cases are not suitable for inclusion in a study of the relation between compensatory and punitive awards. In some cases, the punitive award was capped due to a state or federal cap on punitive damages.[16] Courts in such cases are more constrained than in a case without statutory caps. Several cases without statutory caps on punitive damages also are not usable. In some cases, punitive awards with only nominal compensatory awards were revised or struck down.[17] Even when nominal damages are approved,[18] the punitive-compensatory ratio in such cases is too extreme to be useful for analysis. In other cases, courts approved a

[15] Theodore Eisenberg, John Goerdt, Brian Ostrom, and David Rottman, *Litigation Outcomes in State and Federal Courts: A Statistical Portrait*, 19 Seattle L Rev 433, 434 (1996).

[16] *Luciano v Olsten Corp.*, 110 F3d 210 (2d Cir 1997); *Emmel v Coca-Cola Bottling Co. of Chicago*, 95 F3d 627 (7th Cir 1996); *Greenway v Buffalo Hilton Hotel*, 951 F Supp 1039 (WDNY), aff'd as modified, 143 F3d 47 (2d Cir 1998); *Mobil Oil Corp. v Ellender*, 934 SW2d 439 (Tex App–Beaumont), aff'd in part, rev'd in part, 968 SW2d 917 (Tex 1998); *Jonasson v Lutheran Child and Family Servs.*, 115 F3d 436 (7th Cir 1997); *Benson v Northwest Airlines, Inc.*, 1997 WL 122897 (D Minn 1997); *Hearn v General Electric Co.*, 927 F Supp 1486 (MD Ala 1996); *Anderson v YARP Restaurant, Inc.*, 1997 WL 27043 (SDNY 1997) (but Title VII cap not reached); *Iannone v Frederic R. Harris, Inc.*, 941 F Supp 403 (SDNY 1996) (but Title VII cap not reached); *Williams v Pharmacia, Inc.*, 956 F Supp 1457 (ND Ind), aff'd, 137 F3d 944 (7th Cir 1998).

[17] *Smallwood v Fisk*, 934 P2d 557 (Or App 1997); *Lee v Edwards*, 101 F3d 805 (2d Cir 1996) ($1 nominal damages, $200,000 punitive award ordered remitted to $75,000); *Creative Demos, Inc. v Wal-Mart Stores, Inc.*, 955 F Supp 1032 (SD Ind) ($137 compensatory award, $6,500,000 punitive award struck down), vacated, 142 F3d 367 (7th Cir 1998); Bain v City of Springfield, 678 NE2d 155 (Mass 1997) (zero compensatory, $100,000 punitive award remanded for new trial); Southeastern Sec. Ins. Co. v. Hotle, 473 SE2d 256 (Ga App 1996) (nominal compensatory damages).

[18] See, for example, *Williams v Brimeyer*, 116 F3d 351 (8th Cir 1997). We considered compensatory damage of less than $500 to be nominal.

punitive award even in the absence of a compensatory award.[19] In some cases, appellate courts held that no punitive award at all was warranted.[20] Such cases cannot provide meaningful information about the relative size of the punitive and compensatory award, other than information that the punitive award should be zero. In other cases, the appellate court recast the merits of the case or changed the compensatory award.[21] Cases with these characteristics cannot be used to study the relationship between a punitive award and a non-nominal compensatory award.

Applying these filters yields 251 cases with usable compensatory and punitive awards that were considered by a court, at least in part, for the relation between the two awards.[22] The 251 cases include 165 state court cases and 86 federal cases. All but one of the state cases were appellate court decisions.[23] Of the 86 federal cases, 45 were district court opinions reviewing or discussing a punitive award and 41 were appellate opinions reviewing or discussing a punitive judgment entered by the district court. Of the 251 cases, 114 (45 percent) were decided before *BMW* and 137 were decided after. State cases comprised about 62 percent of the post-*BMW* portion of the sample and about 71 percent of the pre-*BMW* portion of the sample. Little seems to turn on

[19] See, for example, *Harris v Chapman*, 97 F3d 499 (11th Cir), cert. denied, 117 S Ct 2422 (1997); *Jacque v Steenberg Homes, Inc.*, 563 NW2d 154 (Wis 1997) ($1 compensatory, $100,000 punitive award reinstated); *Peter Scalamandre & Sons, Inc. v Kaufman*, 113 F3d 556, 564 & n6 (5th Cir 1997) (noting Texas law prohibited punitive awards without a compensatory award but that the law was later amended); *Shea v Galaxie Lumber & Constr. Co.*, 1997 WL 51655 (ND Ill) ($1 compensatory, $2,500 punitive sustained), rev'd, 1998 WL 498591 (7th Cir 1998); *Sheffield v Andrews*, 679 So 2d 1052 (Ala) ($10 compensatory, $1,000,000 punitive sustained), cert. denied, 117 S Ct 610 (1996).

[20] See, for example, *Langmead v Admiral Cruises, Inc.*, 696 So 2d 1189 (Fla App 3 Dist 1997); *Charles Shaid of Pennsylvania, Inc. v George Hyman Constr. Co.*, 947 F Supp 844 (ED Pa 1996); *Kunewa v Joshua*, 924 P2d 559 (Hi App 1996).

[21] See, for example, *Annis v County of Westchester*, 939 F Supp 1115 (SDNY), aff'd in part, vacated in part, 136 F3d 239 (2d Cir 1998); *Green Bay Packaging, Inc. v Preferred Packaging, Inc.*, 932 P2d 1091 (Okla 1996); *BE & K Constr. Co. v United Brotherhood of Carpenters and Joiners of America*, 90 F3d 1318 (8th Cir 1996); *SK Hand Tool Corp. v Dresser Inds., Inc.*, 672 NE2d 341 (Ill App 1 Dist. 1996); *Molenaar v United Cattle Co.*, 553 NW2d 424 (Minn App 1996); *Call v Heard*, 925 SW2d 840 (Mo en banc), cert. denied, 117 S Ct 770 (1997).

[22] A few cases contained two different compensatory and punitive awards. In these cases we combined the awards and treated the case as a single observation. This differs from our treatment of two such multiple awards cases in a prior study. Eisenberg & Wells, *After BMW* at 409 (cited in note 4). One case required a present value calculation to assess the relationship between the punitive and compensatory awards. *Ingalls v Paul Revere Life Ins. Group*, 561 NW2d 273 (1997). We used a six percent discount rate.

[23] We exclude three opinions published by the City Court of Yonkers in small landlord-tenant disputes.

the precise mix of federal and state cases. To help put the instant sample in perspective, the study based on the *BMW* search included 42 cases decided after *BMW*.[24] The trial level study included 173 punitive damages cases decided during fiscal 1991-1992.[25]

The filtered nature of the sample[26] limits the conclusions that can be drawn. For example, the sample does not permit studying the propriety or frequency of striking a punitive damages award on the ground that no punitive damages were warranted. Cases in which punitive damages are entirely inappropriate are not useful for studying the relation between punitive and compensatory awards. Such cases, however important in their own right, are not in our sample. We are interested in the relation between the compensatory and punitive award conditional on the propriety of both a compensatory and a punitive award.

C. Descriptive Statistics

Table 1 presents summary statistics of key variables in the study. It shows a mean punitive damages award of $6.8 million, and a median of $200,000. Compensatory awards have a mean of about $4.0 million and a median of $87,000. The mean ratio between compensatory and punitive damages is 37.2, but the median is a much smaller 1.8. Large awards skew the distributions of means for punitive and compensatory awards, as well as their ratio. The logarithms of these variables are more amenable to statistical analysis. For some purposes, we will work with the logarithms, which are also summarized in Table 1. Table 1 also shows that federal court cases account for 34 percent of the sample.[27] Slightly more than half the cases were decided after *BMW*.[28]

[24] One of the 42 cases that resulted in an opinion published after *BMW* was decided based on a decision rendered before *BMW* was decided. We have reclassified that case as being pre-*BMW*.

[25] Eisenberg et al, *Predictability of Punitive Damages* at 637 (cited in note 5).

[26] Although the sample is limited to cases with ascertainable compensatory and punitive awards, its size is consistent with a Washington Legal Foundation study counting 284 punitive damages verdicts appealed to state and federal courts in 1993. See Thomas Koenig, *The Shadow Effect of Punitive Damages: Their Effect on Bargaining, Litigation, and Corporate Behavior*, 1998 Wis L Rev 169, 207 ("Koenig, *Shadow Effect of Punitive Damages*"). Many awards were reversed in toto, id, which would make then unsuitable for a study like this one.

[27] The Washington Legal Foundation found one-quarter of the punitive cases appealed in 1993 to be federal. Id at 207.

[28] Interestingly, judges or arbitrators account for 26% of the punitive damages awards, confirming the surprising prominence of non-jury adjudicators in punitive damages cases. Eisenberg & Wells, *After BMW* at 416 n71 (cited in note 4). In at least

Table 1. Descriptive Statistics, Punitive Damages and Other Case Characteristics

	mean	*median*	*standard deviation*	*minimum*	*maximum*	*n*
Punitive Damages (thousands)	6,779	200	75,957	1.00	1,200,000	251
Compensatory Damages (thousands)	4,021	87	47,616	0.50	750,000	251
Ratio-punitive/compensatory	37.21	1.81	279.51	0.01	3,750	251
Punitive Damages (log)	12.24	12.21	2.29	6.91	20.91	251
Compensatory Damages (log)	11.52	11.37	2.14	6.21	20.44	251
Ratio-log punitive/log compensatory	1.08	1.05	0.19	0.61	2.19	251
Federal Case Dummy Variable	0.34	0.00	0.48	0.00	1.00	251
Post-*BMW* Case Dummy Variable	0.55	1.00	0.50	0.00	1.00	251

Source: Westlaw Allfeds and Allstates data bases

Appendix Table 1 reports the breakdown of cases by case category. The results are consistent with prior findings about the kinds of cases that dominate punitive damages awards. Fraud, employment-related cases, and insurer misconduct cases are the most frequent case categories,[29] accounting for about 42% of the cases. Two visible areas of tort law, products liability and medical malpractice, account for only 4.4% of the sample. Including asbestos cases as products cases would increase that number to about 6.4%. These low numbers are consistent with other studies of products liability and medical malpractice cases.[30]

Panel C of the same table shows that the median punitive-compensatory ratios are highest in insurer misbehavior cases. For case categories with more than ten punitive awards insurer misbehavior cases also have the highest median punitive awards. Their compensatory awards are not so extreme. So the high median ratio in insurer misbehavior cases results more from high punitive awards than from low compensatory awards. High awards against insurers probably arise from their large assets and from the decisionmakers' outrage at these sophisticated companies' efforts to fail to honor their obligations.[31]

one state, the decision to award punitive damages can be made by a jury but the amount of the award is determined by a judge. The results reported here are not dependent on the presence of non-jury awards in the sample.

[29] Eisenberg et al, *Predictability of Punitive Damages* at 635 (cited in note 5).

[30] See id at 635-36.

[31] See, for example, *Walston v Monumental Life Ins. Co.*, 923 P2d 456 (Id 1996); *Cock-N-Bull Steak House, Inc. v Generali Ins. Co.*, 466 SE2d 727 (SC 1996).

Appendix Table 2, which breaks down results by jurisdictions, shows that Alabama and Texas are the only two states that comprise more than five percent of the sample. But the most extreme punitive-compensatory ratio is in California and not in either Texas or Alabama. California's high ratios in this database are consistent with a large RAND trial level data set for California.[32] Alabama's lack of distinction suggests that its reputation for high punitive awards may stem from high compensatory awards leading to high punitive awards and not simply from dramatically inflated punitive awards at constant levels of compensatory awards.

Significance levels reported in the Appendix shed further light on a possible Alabama effect. The interstate differences in punitive and compensatory damages are significant or nearly significant based on a Kruskal-Wallis test. But the interstate differences in the punitive-compensatory ratios are not statistically significant. These results hold if one excludes federal cases from the analysis. In short, Alabama's reputation may stem in part from focusing on high punitive awards without examining the punitive-compensatory ratio. In addition, Alabama has a higher percentage of insurer misbehavior cases than any other state in this study.[33] Since insurer misbehavior cases generate high punitive-compensatory ratios, another source of any Alabama effect may be its many insurer misbehavior cases. But the high punitive-compensatory ratio in insurance cases is not solely attributable to Alabama. Excluding Alabama cases, the median ratio is still a high 5.3.

Case category and interstate effects summarized in the Appendix are not significant in explaining median punitive-compensatory ratios. In the following analysis, we do not report case category or interstate results. None of our principal results would differ materially if we included them in the analysis.

II. PUNITIVE DAMAGES AWARDS IN PUBLISHED OPINION CASES

Figure 1 presents, inter alia, a scatterplot of the relation between punitive and compensatory awards in the published opinion sample. The scatterplot suggests a strong correlation between the logarithm of punitive awards and the logarithm of compensatory awards. To quantify that relation, we use robust regression to model the loga-

[32] See Eisenberg & Wells, *After BMW* at 403-06 (cited in note 4).

[33] Five of the (22%) 23 insurer misbehavior cases in the sample are from Alabama. Yet Alabama accounts for less than 10% of the cases in the sample. See Appendix Table 2.

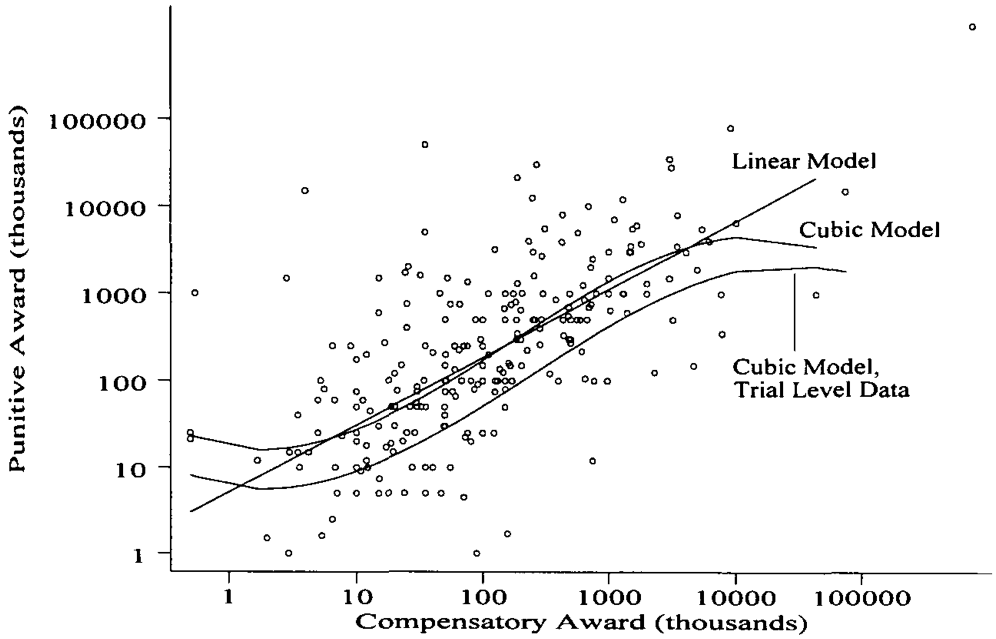

Figure 1. Punitive and Compensatory Awards (log scales)—Punitive and compensatory awards, cubic robust regression model, and linear robust regression model, all based on published opinions, mid-1995 to mid-1997, and cubic regression model based on CTCN trial level data, 1991-1992. Sources: Westlaw Allfeds and Allstates databases; CTCN, 1991-1992, Inter-University Consortium for Political and Social Research, Civil Justice Survey of State Courts, 1992: [United States] (ICPSR 6587).

rithm of punitive damages as a function of the logarithm of compensatory damages.[34] The simple robust regression model yields the following equation, in logs:

$$\text{punitive award} = .779 \times \text{compensatory award} + 3.164 \qquad (1)$$

More detailed model specifications, including dummy variables for case categories and jurisdictions, yield a similar coefficient for the compensatory award variable. Figure 1 includes the straight line defined by equation (1), labeled as "Linear Model."

Theoretical considerations and prior empirical results suggest fitting a cubic model to the data. The punitive-compensatory multiple ought to be harder to pinpoint for low than for high compensatory

[34] Robust regression allows us to deal with the outliers in the data. Classical least squares is robustized by, instead of minimizing a sum of squares, minimizing a sum of a less rapidly increasing function. Peter J. Huber, *Robust Statistics* (Wiley, 1981). The net effect is that outliers are down weighted. Various robust estimators were applied to this regression model and all yielded essentially identical estimates. The use of log transforms is discussed in Part II.C below.

damages awards. Factfinders outraged enough to award punitive damages in the face of low compensatory awards might be expected to employ higher or less predictable multiples of compensatory awards. This pattern was observed in the trial level data set.[35] In fact, the top curved line in Figure 1 labeled "Cubic Model" provides a better visual fit to the data, at least up to extremely high compensatory awards, and a modestly better explanatory model than the simple linear model.[36] Figure 1 also shows that, for a large range of compensatory awards, from about $10,000 to about $10 million, the linear and cubic models are not materially different.

The strong, significant correlation between punitive and compensatory damages invites comparison with prior studies finding a similar relation. Table 2 provides data to facilitate comparing the instant results to other punitive damages studies. Panel A reports the mean and median punitive award and the mean and median punitive-compensatory ratio in several data sets. Panel B reports the best-fitting linear robust regression line in several data sets.[37] We first compare this study's results with a previous study of published opinions. We then compare the results with a previous study of trial level outcomes.

A. Comparison With Prior Study of Published Opinions

We previously compared the pattern of punitive and compensatory awards in the cases generated by the *BMW* search with the pattern

[35] See Eisenberg et al, *Predictability of Punitive Damages* at 657 (Table 7) (cited in note 5).

[36] As before, the adjusted r-squared of a cubic model slightly exceeds that of a linear model. Eisenberg & Wells, *After BMW* at 392 (cited in note 4). A kernel smoothed graph of punitive damages (log) as a function of compensatory damages (log) yields a line that approximates a cubic. The kernel smoothed curve gives a nonparametric estimate of the true relationship between punitive and compensatory damages and does not depend on any a priori specified functional form (such as linear, polynomial, or exponential). See generally M.P. Wand & C.M. Jones, *Kernel Smoothing* (Chapman and Hall, 1995).

[37] In models using this study's data that include a dummy variable to represent whether a state caps punitive damages, the coefficient for the dummy variable is negative but statistically insignificant. Nor is it significant in models using the CTCN trial level data. We treat Colorado, Florida, Illinois, Indiana, Kansas, North Carolina, North Dakota, and Virginia as having broad-based caps on punitive awards. Colo Rev Stat § 13-21-102(1)(a) (1988); Fl Stat ch. 768.73(1)(a) (1995); 735 Ill Comp Stat Ann 5/2-115.05 (West 1993); Ind Code Civ Proc 34-51-3-4 (1998); Kan Stat Ann § 60-3701(e) (West Supp 1996); NC Gen Stat § 1D-25 (1997); ND Cent Code § 32-03.2-1(4) (1996); Va Code Ann § 8.01-38.1 (Michie 1992). We treat Georgia, Nevada, and New Jersey as having caps applicable to some but not all case categories. Ga Code Ann § 51-12-5.1(g) (Supp 1996); Nev Rev Stat § 42.005 (1995); NJ Stat Ann 2A:15-5.14 (1998).

Table 2. Punitive Damages Awards by Case Sample

A. Mean and Median Punitive Award and Punitive-Compensatory Ratio by Case Sample

case sample	*punitive award (thousands)*		*punitive-compensatory ratio*		*number of cases*
	mean	*median*	*mean*	*median*	
federal appeals (*BMW* search)	154,639	750	35.3	6.2	8
federal district court (*BMW* search)	2,312	688	8.9	7.9	12
state appellate court (*BMW* search)	4,419	937	241.6	13.6	22
state trial court (CTCN)	544	50	14.9	1.3	173
this study of published opinions	**6,779**	**200**	**37.2**	**1.8**	**251**

Sources: Eisenberg & Wells, 1998 Wisc. L. Rev 387; Westlaw Allstates and Allfeds databases

B. Robust Regression Results, Four Data Sets

dependent variable = logarithm of punitive damages

Data set	*coefficient for logarithm of compensatory damages*	*constant*	*n*
published opinions, this study	.779 (.047)	3.164 (.550)	251
CTCN 1991-92	.782 (.059)	2.094 (.679)	173
Rand Cook County 1960-84	.819 (.044)	1.879 (.478)	208
Rand San Francisco 1960-84	.810 (.033)	2.382 (.366)	616

standard errors in parentheses

Sources: Eisenberg et al., 26 J. Leg. Studies 623; Westlaw Allstates and Allfeds databases.

of awards in the trial level sample.[38] The comparison suggested that describing or assessing punitive awards based on the reported opinions in that study overstates the level of punitive awards and the ratio of punitive awards to compensatory awards. This results because published opinions are a filtered set of cases that often does not represent the mass of cases.[39]

[38] Eisenberg & Wells, *After BMW* at 414-15 (cited in note 4).

[39] See, for example, Theodore Eisenberg & Sheri Lynn Johnson, *The Effects of Intent: Do We Know How Legal Standards Work?*, 76 Cornell L Rev 1151, 1172-75 (1991); Theodore Eisenberg & Stewart J. Schwab, *What Shapes Perceptions of the Federal Court System?*, 56 U Chi L Rev 501 (1989).

Panel A confirms this finding using a more complete set of published opinions. It includes both pre- and post-*BMW* opinions, and the post-*BMW* opinions are not limited to those in which a citation to *BMW v. Gore* appears. This more complete published opinion sample confirms that the means, medians, and ratios are higher in published opinions than in trial level cases. But panel A suggests that the published opinion bias, while substantial, is less dramatic in a fuller sample of published opinions than in the published opinions resulting from the *BMW* search. The median punitive award and the median punitive-compensatory ratio in the fuller sample of published opinions are much closer to those in the national trial level sample. The median state trial court punitive award is less than one-tenth the median punitive award in opinions resulting from the *BMW* search. But the median state trial court punitive award is 25 percent of the median award in the fuller published opinion sample, which includes the *BMW*-based cases. The median punitive-compensatory ratio in state trial court cases is about one-sixth the median ratio of *BMW*-based cases. The trial court ratio is about 72 percent of the punitive-compensatory ratio in the fuller published opinion sample.

B. Comparison with Trial Level Study

The lower curved line in Figure 1, labeled "Cubic Model, Trial Level Data," is the cubic model from the CTCN trial level sample. It is strikingly similar to the cubic model in the published opinion data set. Only the intercept materially differs.[40]

Panel B's first two rows report the robust regression results from equation (1) for our published opinion sample. The next three lines are from our earlier work with trial data.[41] Together, these results show that the slope for published opinion cases is close to the slopes in the CTCN and RAND trial level data sets. All slopes lie within each other's 95% confidence intervals and all are highly statistically

[40] For the CTCN trial level sample, the equation that bests fits the data is:

$$y = -.013x^3 + .498x^2 - 5.220x + 25.434,$$

where y is the logarithm of the punitive award and x is the logarithm of the compensatory award. Eisenberg & Wells, *After BMW* at 392, 393 n13 (cited in note 4). For the instant study's published opinion sample, the cubic equation that best fits the data is:

$$y = -.016x^3 + .560x^2 - 5.771x + 28.048,$$

This equation is based on robust regression. See Note, *Shadow of BMW* (cited in note 4).

[41] See Eisenberg et al, *Predictability of Punitive Damages* at 651 (cited in note 5).

significant. We cannot reject the hypothesis that the line with the same slope best models the relationship between compensatory and punitive damages in all four data sets. But the constant for this study's published opinion data differs more noticeably from the constants in the trial level data sets. Although published opinion cases have higher stakes than trial level cases, the relation between compensatory and punitive awards is largely invariant across case samples.

As before,[42] the difference between the mass of cases decided in state trial courts and the filtered set of cases that lead to reported opinions has implications for judicial review of the ratio of punitive to compensatory awards. Courts reviewing punitive awards often try to assess such awards by comparing them with prior decisions.[43] Courts that search for prior patterns of awards can be expected to find only cases reporting punitive awards in cases with reported opinions, such as those reported by Westlaw. Yet comparing punitive-compensatory ratios with the biased sample of reported opinions elevates the ratios that appear to have been approved in the past.

C. The Persistence of the Log-Log Relation Between Compensatory and Punitive Awards

Questions have been raised about the log transform of dollar amounts needed to reveal the relation between compensatory and punitive awards. Sunstein et al. suggest concern about that relation in noting that "defendants and plaintiffs live in a world of real dollars, not of log dollars."[44] Viscusi states that "[t]aking logs of award amount compresses the extent of the variation in what is of actual concern to defendants, which is the level of punitive damages, not its log."[45]

[42] Eisenberg & Wells, *After BMW* at 413-16 (cited in note 4).

[43] See, for example, *Bain v City of Springfield*, 678 NE2d 155, 162 (Mass 1997); *Williams v ITT Financial Servs.*, 1997 WL 346137 (Ohio App 1 Dist), at *19, appeal allowed, 685 NE2d 545 (Ohio 1997).

[44] Cass R. Sunstein, Daniel Kahenman, and David Schkade, *Assessing Punitive Damages (With Notes on Cognition and Valuation in Law)*, 107 Yale LJ 2071, 2153 n21 (1998) ("Sunstein et al, *Assessing Punitive Damages*").

[45] W. Kip Viscusi, *Why There is No Defense of Punitive Damages*, 87 Geo LJ 381, 386 (1998) ("Viscusi, *No Defense*"). Viscusi also seems to treat our reporting of a surprisingly regular relation between compensatory and punitive damages as somehow claiming that our data show that punitive damages have a deterrent effect. Id at 385. That is not an issue we address. We note that Viscusi's showing that punitive damages do not deter rests on an erroneous view of states' punitive damages regimes. Theodore Eisenberg, *Measuring the Deterrent Effect of Punitive Damages*, 87 Geo LJ 347, 348-49 (1998). This mistake is not fully corrected in Viscusi's subsequent analysis. See Viscusi, *No Defense* at 391-92 (trying to account correctly for Louisiana

Two issues relating to the log-transform are worth separating. One is that the relation between compensatory and punitive awards only emerges if awards are transformed into logarithms.[46] This concern is of questionable validity. Many important economic relations only emerge, or emerge most clearly, when amounts are transformed into logs. For example, the Nobel-prize-winning Black-Scholes pricing of options and corporate liabilities assumes that the natural logarithms of stock prices are normally distributed.[47] Hence, in their model, real world prices are measured in a log dollar world. Financial officers of large companies who obtain advice about markets do not ignore or denigrate the model because it depends on a log transform of dollar amounts. These same financial officers, and their risk-management colleagues, cannot credibly question the validity of other financial effects that emerge only after log-transforms.

A second concern, also voiced by Sunstein et al., and echoed by Viscusi,[48] is that a given compensatory award can support a wide range of punitive awards and that the log transform suppresses this variability. Sunstein et al. state:

> In terms of real dollars, the judgments of our respondents and of the juries examined by Eisenberg et al. are correctly described

but ignoring his possibly erroneous characterization of Michigan). Indeed, if one recharacterizes Michigan as a state in which businesses must worry about possible exemplary damages, it appears that Viscusi's own data show that both medical misadventure death rates and total accident death rates are lower in states with punitive damages than in states without them. Id at 392 (Table 1). Furthermore, recharacterizing Michigan would lead to a higher rate of toxic chemical accidents involving injury or death in states without punitive damages than in states with them. See W. Kip Viscusi, *The Social Costs of Punitive Damages Against Corporations in Environmental and Safety Torts*, 87 Geo LJ 285, 290 (Table 1) (1998).

[46] Examining the raw punitive damages data reveals that the data are positive and skewed left. This shape is quite common when dealing with data measured in dollars. A distribution that fits such data is the log normal distribution, that is, the natural logarithm of the data has a normal distribution. Therefore it is natural to model the natural logarithm of punitive damages because their distribution is essentially normal. When constructing a regression model relating punitive and compensatory damages, it is reasonable to measure the two variables on the same scale. Hence a log-log type regression model accounts for the skewness of punitive damages and also measures the most important predictor variable on the same scale. Using a model other than the log-log link function would violate the important normality assumption in regression analysis and give misleading regression results.

[47] Fischer Black & Myron Scholes, *The Pricing of Options and Corporate Liabilities*, 81 J Pol Econ 637 (1973). See also Edward J. McCaffery, Daniel I. Kahneman & Matthew L. Spitzer, *Framing the Jury: Cognitive Perspectives on Pain and Suffering Awards*, 81 Va L Rev 1341, 1358-59, 1368 (1995) (study by one of Sunstein's co-authors using log transformation of dollar amounts).

[48] Viscusi, *No Defense* at 386 (cited in note 45).

> as erratic and unpredictable, because the severe skewedness creates the possibility of either modest or disastrous losses in identical cases. This produces unfairness and could induce risk aversion even in very large firms.[49]

The "severe skewedness," of course, is precisely what the log transform eliminates, as Sunstein et al. acknowledge. So if one were to think in terms of log awards, the range of awards is distinctly less erratic and unpredictable. We do not deny that people tend to think in untransformed dollars. That they do so does not undermine the power of a model, and its predictive abilities, in transformed dollars. Most importantly, the logarithmic transformation is monotonic, so that a significant relation between log punitive damages and log compensatory damages implies a significant relation between punitive damages and compensatory damages. Remaining concerns about the possibility of disastrous losses could be addressed by a cap of ten times compensatory damages, as discussed below.

III. PUNITIVE DAMAGES AWARDS BEFORE AND AFTER BMW

It is natural to expect the pattern of punitive damages awards to change after *BMW* applied constitutional limitations to reduce a punitive award. And it has been asserted that *BMW* did substantially change review of punitive awards. As one commentator noted, "courts throughout the country . . . have found the *BMW* guideposts to provide a useful analytical framework and, applying that result, have overturned numerous verdicts on excessiveness grounds."[50] But it is also possible that *BMW* affects few cases.[51] Punitive awards so high as to be unconstitutional ought to be rare.[52] It is therefore appropriate to test, via something more than anecdotal reporting, whether *BMW* led to a noticeable shift in punitive damages award patterns. This testing is limited by the selection of cases for adjudication. It is possible that *BMW*'s greatest effect is in an unobserved mass of disputes that never reach trial or result in a punitive award.[53] Nevertheless, it seems plausible to expect *BMW* to have had some

[49] Sunstein et al, *Assessing Punitive Damages* at 2153 n21 (cited in note 44).

[50] Trager, *One Year Later* (cited in note 3). To similar effect, see Note, *Shadow of BMW* at 456 (cited in note 4).

[51] Eisenberg & Wells, *After BMW* at 417 (cited in note 4).

[52] Id

[53] For efforts to test such effects, see Koenig, *Shadow Effect of Punitive Damages* (cited in note 26); Herbert M. Kritzer & Frances Kahn Zehmans, *The Shadow of Punitives: An Unsuccessful Effort to Bring It into View*, 1998 Wis L Rev 157.

effect, at least in the short term, on the treatment of punitive damages awards.[54]

A. The Patterns of Awards

Table 3 reports descriptive statistics for the 1995-1997 published opinion cases, broken down by whether the case was decided before or after *BMW*. Some noteworthy differences exist before and after *BMW*. The means, medians, and standard deviations for punitive and compensatory awards increased after *BMW*. But only the standard deviation increase is statistically significant. The median punitive-compensatory ratio, however, decreased and that decrease is statistically significant or nearly significant for both the untransformed ratios (p = .063) and the log transformed ratios (p = .034).

Table 3 contains an important lesson about reporting punitive damages award trends. If one examines only punitive damages awards, Table 3 supports reporting a post-*BMW* increase. The pre-*BMW* mean was about $1.4 million and the post-*BMW* mean was about $11.2 million. The medians and standard deviations show similar patterns of increase. But the median ratio of punitive to

Table 3. Punitive Damages and Punitive-Compensatory Ratios by Pre- or Post-*BMW* Status

	mean		*median*		*Standard deviation*	
	pre-BMW	*post-BMW*	*pre-BMW*	*post-BMW*	*pre-BMW*	*post-BMW*
punitive (thousands)	1,413	11,200	155	218	5,157	102,700*
compensatory (thousands)	406	7,029	75	91	1,001	644,000*
ratio	34.1	39.8	2.15	1.41	217.1	323.2
log punitive	12.11	12.34	11.95	12.29	2.04	2.48*
log compensatory	11.27	11.72	11.23	11.42	1.95	2.27
ratio of logs	1.09	1.07	1.07	1.02	.19	.19
N	114	137	114	137	114	137

Source: Westlaw Allfeds and Allstates databases.
* significant difference between pre- and post-*BMW* cases

[54] See Kevin M. Clermont & Theodore Eisenberg, *Do Case Outcomes Really Reveal Anything About the Legal System? Win Rates and Removal Jurisdiction*, 83 Cornell L Rev 581 (1998) (suggesting that cautious interpretation of data about case outcomes can yield insights into the legal system's operation).

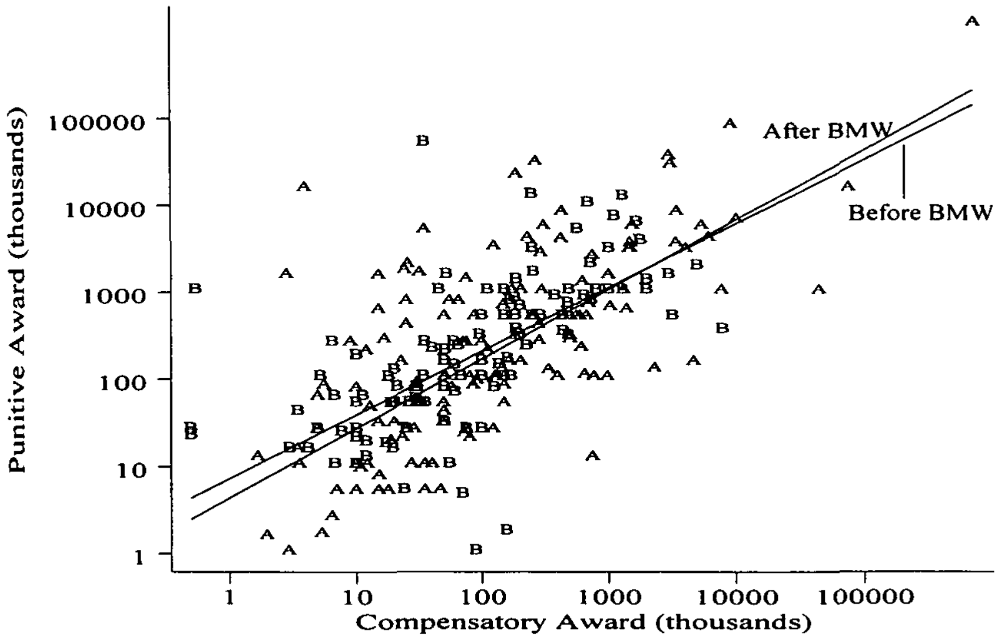

Figure 2. Punitive and Compensatory Awards, Before and After *BMW v. Gore* (log scales)—Punitive and compensatory awards, before and after *BMW v. Gore*, and linear robust regression models, before and after *BMW v. Gore*. Source: Westlaw Allfeds and Allstates databases, mid 1995 to mid 1997.

compensatory awards decreased after *BMW* and the median decrease is nearly statistically significant. If punitive awards should increase because increased harm, as measured by the compensatory award, justifies increased punishment, reporting only increasing punitive damages awards is misleading.

To further explore *BMW*'s effect, we present another scatterplot of punitive and compensatory awards. Unlike the prior scatterplot, however, Figure 2 distinguishes between pre- and post-*BMW* awards. Figure 2 plots the same cases as Figure 1. But it labels cases decided before *BMW* with a "B" and cases decided after *BMW* with an "A."

Figure 2 also contains two straight lines to facilitate comparing pre-and post-*BMW* relations between compensatory and punitive awards. The lines result from the same robust regression model as is reported in equation (1), except that the model is estimated twice, once each for pre- and post-*BMW* awards. This yields a separate best-fitting line for pre- and post-*BMW* awards. Visually, the lines differ slightly in slope and slightly more in intercept. But they do not appear to be materially different. Table 4 reports the results of the two robust regressions. A robust regression model not reported here

Table 4. Robust Regression Results, Before and After *BMW*

	dependent variable = logarithm of punitive damages		
Data set	*coefficient for logarithm of compensatory damages*	*constant*	*n*
this study, pre-*BMW* cases	.732 (.059)	3.821 (.678)	114
this study, post-*BMW* cases	.800 (.069)	2.843 (.822)	137
standard errors in parentheses			

Source: Westlaw Allfeds and Allstates databases.

with an interaction term for post-*BMW* status and compensatory awards, and a post-*BMW* dummy variable, confirms that we cannot reject the hypothesis that the two lines are identical.

The decline in the punitive-compensatory ratio raises the possibility that *BMW* has had the expected effect of reducing the ratio. We further explore this question via Figure 3, which presents kernel density estimates of the distributions of the punitive-compensatory ratios before and after *BMW* was decided. Figure 3 shows a definite shift in the distribution toward lower ratios after *BMW*. A Kolmogorov-Smirnov test of the equality of the two distributions allows rejection of the null hypothesis of equality ($p = .003$). Thus, while tests of the central tendencies of the pre- and post-*BMW* distributions of ratios yielded marginally significant results, both Figure 3 and a test of the entire distributions suggest a notable shift towards lower ratios after *BMW*.

The shift in ratios does not, however, occur in the expected part of the distribution. Although the low ratio awards are more dominant after *BMW*, the high ratio awards have, if anything, also increased. The overall effect is a flatter distribution with a smaller middle and a shift toward both low and high extremes. Although the median punitive-compensatory ratio is lower after *BMW*, the 75th, 90th, and 95th percentiles are all higher than before. Thus, while the distributions differ, it is not clear that *BMW* should be regarded as the cause. *BMW* would seemingly forecast a shift away from the right tail, not growth in it. Regression analysis confirms the story suggested by Figure 3. In models with the punitive-compensatory ratio as the dependent variable, and the compensatory award and a post-*BMW* dummy variable as the independent variables, we find no significant post-*BMW* effect.

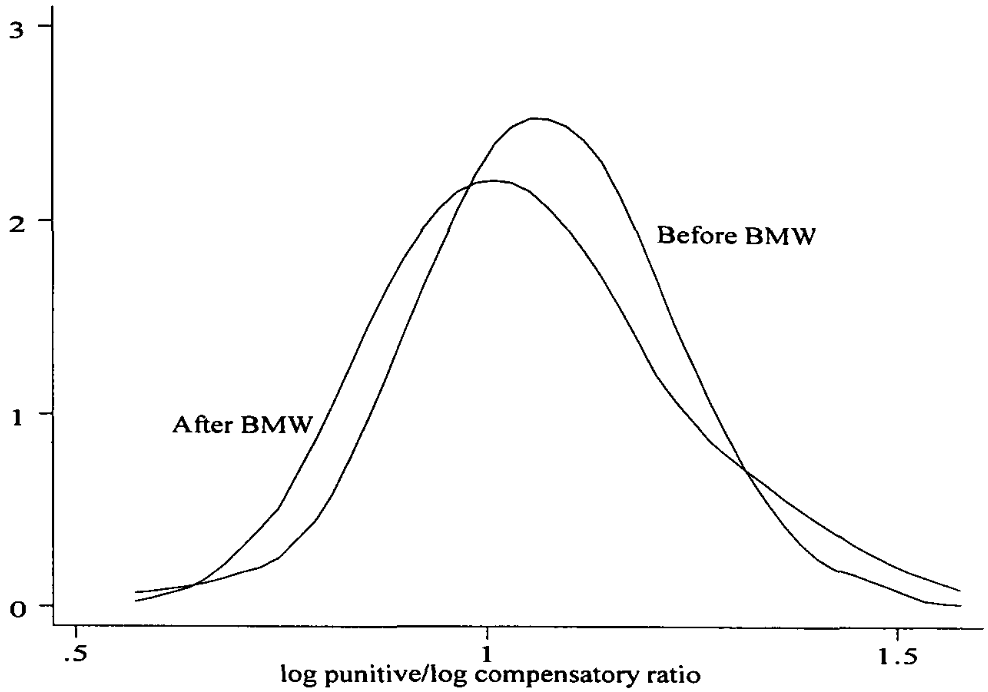

Figure 3. Punitive-Compensatory Ratio (logs), Before and After *BMW v. Gore*—Kernel density estimates of punitive-compensatory ratio distributions before and after *BMW v. Gore.* Source: Westlaw Allfeds and Allstates databases, mid 1995 to mid 1997.

B. The Patterns of Punitive Awards Reduced as Excessive

We first note the rate at which punitive awards were reduced in published opinions before and after *BMW*. Before *BMW*, 13 of 114 awards were reduced as excessive (11.4%). After *BMW*, 18 of 137 awards (13.1%) were reduced. This increase is not statistically significant ($p = .71$). Both reduction rates are quite low,[55] reflecting the fact that most punitive awards are not of the multimillion dollar, headline-grabbing variety. Even in the published opinion database, which provides an upwardly biased sample of punitive awards, only 25% of the punitive awards exceed one million dollars and only 10% exceed $3.9 million. Most punitive awards are not reduced because they are not extremely high.

A word of caution is in order about the statistical insignificance of the difference in pre- and post-*BMW* reduction rates. The power

[55] A Washington Legal Foundation study of punitive damages appeals also found that punitive damages awards were rarely reduced as being excessive. Koenig, *Shadow Effect of Punitive Damages* at 207 (cited in note 26).

of a statistical test is the likelihood of detecting an effect of a specified size at a specified significance level. If a test is not very powerful, the likelihood of detecting the effect is small. Perfectly executed studies may fail to reveal socially important differences "simply because the sample sizes are too small to give the procedure enough power to detect the effect."[56] It is important to report a statistical test's power when one claims that no significant effect has been detected.

A power calculation requires specifying what is a socially meaningful change in the observed proportion of reduced punitive damages awards. The observed change in rate, from 11% to 13%, represents striking approximately two additional punitive awards in about a year. It does not strike us as socially meaningful. Suppose we, admittedly arbitrarily, specified that we would consider meaningful an increase from 11% to 22% in the number of awards that are reduced as being excessive. This would represent an additional 11 strikes per hundred cases per year. At a significance level of .1, the power of our test would be about .70. That is, given our sample size, we have about a 70% chance of detecting a shift from an 11% to a 22% strike rate. So our caution is that the failure to detect a significant drop in strike rates after *BMW* should be regarded as suggestive, not conclusive.

Moreover, studying reduction rates in isolation is incomplete. Since *BMW* emphasized the relation between punitive and compensatory awards, it is reasonable to explore the reduction rate as a function of the punitive-compensatory ratio. We have used several models in which the dependent variable is whether a punitive award is reduced and the explanatory variables include the ratio of the log of punitive awards to the log of compensatory awards. These models include controls for case categories and other variables. Although the ratio variable is highly significant—higher ratio awards are more likely to be reduced—a post-*BMW* dummy variable is not. We thus still lack evidence that *BMW* altered the rate at which punitive awards are struck.

IV. USING TRIAL LEVEL DATA TO PREDICT WHICH PUNITIVE AWARDS WILL BE REDUCED

Figure 1 and Table 2 show that the punitive-compensatory damages relation is similar across the trial level and published opinion data sets. This raises the possibility of using the trial level data to predict the treatment of punitive awards in published opinions. In fact, the trial level punitive-compensatory awards pattern provides a powerful

[56] Stanton A. Glantz, *Primer of Biostatistics* 178 (4th ed, 1997).

tool for forecasting the treatment of punitive awards in published opinions. Both visual inspection of the data and statistical analysis suggest the ability to predict punitive award reductions in published opinion cases on the basis of the trial level data.

Figure 4 presents another scatterplot of Figure 1's published opinion data with two new features. First, we superimpose on the published opinion scatterplot three curved lines derived from the CTCN trial level data. Figure 4's middle curved line, labeled "Best fitting line, trial level data," is the regression line that best fits the trial level data. The curved lines labeled "Upper 95% prediction line, trial level data" and "Lower 95% prediction line, trial level data," are the upper and lower 95 percent prediction lines for the cubic model of the trial level data. As we have noted, these lines provide a way to designate punitive awards that, given the mass of trial level awards, bear an extreme relation to compensatory awards.[57] Thus, scatterplot points above the upper 95% line are extreme in the sense that they represent punitive-compensatory award relations that are above the trial level data upper 95% prediction band for such relations. Second, Figure 4 represents the individual published opinion

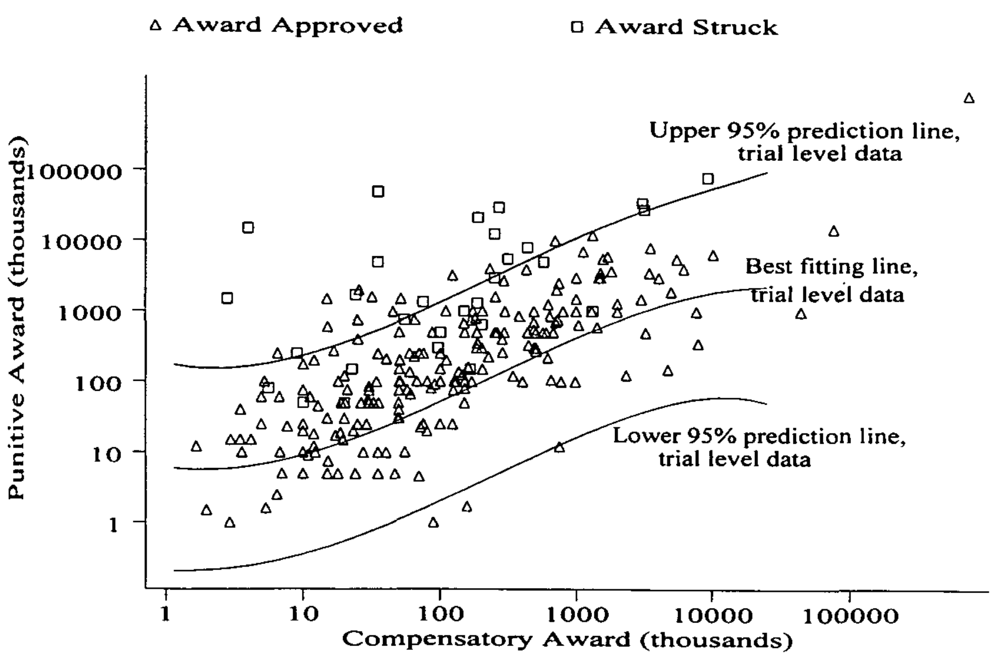

Figure 4. Punitive and Compensatory Awards (log scales)—Upper 95% prediction lines based on trial level data, and punitive and compensatory awards from the published opinion data. Sources: CTCN, 1991-1992, Inter-University Consortium for Political and Social Research, Civil Justice Survey of State Courts, 1992: [United States] (ICPSR 6587); Westlaw Allstates and Allfeds databases.

[57] Eisenberg & Wells, *After BMW* at 410-16 (cited in note 4).

cases by symbols distinguishing between cases in which punitive awards were deemed excessive and cases in which they were not. The figure visually suggests that awards above the trial level 95% prediction line are more likely to be deemed excessive than are awards below that level.

Table 5 explores three methods of predicting whether a punitive award will be reduced as excessive in the published opinion data. The first method relies on Figure 4's upper 95% prediction line. It forecasts that a punitive award will be reduced if the relation between the punitive and compensatory award in a published opinion case falls above the upper 95% prediction line in Figure 4. The second method tested forecasts reduction of an award that exceeds the greater of $300,000 or three times the compensatory award, a popular capping system.[58] The third method forecasts reduction of a punitive award that exceeds the greater of $500,000 or ten times the compensatory award. The ten-times approach is empirically derived from our work with the trial level sample.[59]

Table 5 shows the percent of punitive award treatments in published opinion cases correctly classified by the three methods. Table 5 separately reports how successfully the three methods predict reductions and non-reductions. For example, when the trial level data upper 95% line forecasts that no reduction will occur, Table 5's first "correctly classified" row shows that upper 95% line is correct for 93.7% of the cases. When the line forecasts that a reduction will occur, it is correct in 60.7% of the cases. Since reduction of an award is a rare event, occurring in about 12% of the cases, this 60.7% rate is reasonably impressive. It far exceeds the correct classification reduction rate for the second method, based on three times the compensatory award. The three-times system, reported in Table 5's middle numerical column, correctly forecasts a reduction in award in only 43.4% of the cases. Comparing Table 5's third numerical column to its first column shows that there is not much difference between using the upper 95% prediction line and a simple forecasting system based on the greater of $500,000 or ten times the compensatory award.

Even 90% accuracy in forecasting treatment of punitive awards is in one sense less impressive than it first appears. A naive model that always forecasts no reduction in a punitive award will be correct 87.7% of the time, as indicated in the first column of Table 5's last row. An alternative way to assess forecasting methods is to calculate how much they improve over the naive model. Table 5's last row

[58] Id at 401.
[59] Id at 415.

Table 5. Cases Correctly Predicted Using Three Criteria for Deeming Published Opinion Punitive Awards Excessive

	trial level data cubic model upper 95% line	*greater of $300,000 or 3 times compensatory award*	*greater of $500,000 or 10 times compensatory award*
reduction in punitive award *not* predicted based on column criteria			
correctly classified	93.7	96.0	93.3
incorrectly classified	6.3	4.0	6.7
reduction in punitive award predicted based on column criteria			
correctly classified	60.7	43.4	59.3
incorrectly classified	39.3	56.6	40.7
total correctly classified	90.0	84.9	89.6
% reduction in error over naive model of never reduced (87.7% correct)	19.0	-22.3	15.7

shows that prediction based on the trial level data's upper 95% prediction line reduces forecasting errors by 19.0% over the naive model. The three-times system generates 22% *more* erroneous predictions than the naive model. The ten-times system reduces forecasting errors by 16.4% over the naive model. We conclude that the trial level data provide a useful tool for forecasting when punitive awards will be reduced.

Logistic regression models based on Table 5's column criteria (one criterion in each model) and dummy variables for case categories and states do not materially improve on the simple tabular models reported here.

V. CONCLUSION

Punitive and compensatory damages awards exhibit the same basic relation in published opinion cases and in trial level data. Published opinion cases tend to have higher stakes and, therefore higher awards. But the relation between punitive and compensatory awards is stable.

We find little evidence that *BMW* caused a noticeable shift in the pattern of punitive and compensatory awards. It may be that insufficient time has elapsed for *BMW*'s effect to be fully felt. But it also may be that the case is of less practical importance than some believe. *BMW* is a constitutional decision and one should not expect the Constitution to play a role in routine punitive damages cases. The punitive-compensatory ratio in *BMW* was extraordinary, thereby promoting constitutional intervention. The case provides a direct precedent for policing punitive awards only in similarly extreme cases.

We also find little evidence to support substantial reported regional differences in punitive-compensatory ratios. In our published opinion data, we find no significant regional variation in the punitive-compensatory ratio.

APPENDIX

Table 1. Punitive and Compensatory Awards by Case Category

	mean	*median*	*std. dev.*	*n*
A. Punitive Damages (thousands)				
Assault, Battery	196	75	231.2	10
Attorney Misconduct	670	106	1,277.8	7
Conversion	17	11	16.5	6
Defamation	3,290	700	5,506.2	7
Employment Discrimination	5,985	150	18,370.7	25
Employment Other	516	272	682.7	15
Fraud	1,997	300	4,223.2	43
Insurer Misbehavior	2,399	600	4,077.8	23
Interference with Contractual Relations	4,430	375	10,382.5	8
Police, Prison Authority Misconduct	79	25	141.7	11
Products Liability	6,807	1,000	12,652.5	7
Other	14,487	150	127,132.8	89
Total	*6,778*	*200*	*75,956.8*	*251*
Significance	*0.001*	*0.000*		
B. Compensatory Damages (thousands)				
Assault, Battery	137	84	156.7	10
Attorney Misconduct	346	32	558.3	7
Conversion	31	31	19.7	6
Defamation	714	675	798.2	7
Employment Discrimination	510	50	1,834.7	25
Employment Other	239	100	297.3	15
Fraud	3,157	126	13,284.4	43
Insurer Misbehavior	178	52	302.5	23
Interference with Contractual Relations	1,076	210	1,568.5	8
Police, Prison Authority Misconduct	40	10	73.3	11
Products Liability	4,125	3,048	3,989.7	7
Other	9,058	125	79,443.7	89
Total	*4,021*	*87*	*47,616.1*	*251*
Significance	*0.000*	*0.001*		

C. Punitive-Compensatory Ratio				
Assault, Battery	1.64	1.24	1.5	10
Attorney Misconduct	4.00	1.54	6.1	7
Conversion	1.60	0.57	2.7	6
Defamation	536.93	1.00	1416.8	7
Employment Discrimination	66.91	2.50	285.1	25
Employment Other	6.22	3.33	12.5	15
Fraud	10.90	2.00	23.5	43
Insurer Misbehavior	117.08	6.67	392.1	23
Interference with Contractual Relations	19.67	0.77	39.4	8
Police, Prison Authority Misconduct	7.63	2.00	12.3	11
Products Liability	2.32	1.00	4.1	7
Other	3.85	1.46	9.1	89
Total	*37.21*	*1.81*	*279.5*	*251*
Significance	*0.004*	*0.009*		

APPENDIX

Table 2. Punitive and Compensatory Awards by Jurisdiction

	mean	*median*	*std. dev.*	*n*
A. Punitive Damages (thousands)				
Alabama	2,227	150	4,159.3	17
Arkansas	122	87	157.1	8
California	10,898	1,500	23,213.6	12
Federal	15,616	250	129,378.4	86
Ohio	1,068	25	2,645.4	9
Texas	1,365	800	2,238.6	23
other states	1,541	120	4,545.8	96
Total	*6,779*	*200*	*75,956.8*	*251*
Significance	*0.002*	*0.003*		
B. Compensatory Damages (thousands)				
Alabama	323	57	590.2	17
Arkansas	82	53	88.3	8
California	1,444	308	2,643.2	12
Federal	10,668	75	81,224.9	86
Ohio	491	25	1,156.5	9
Texas	678	185	1,174.4	23
other states	504	78	1,130.4	96
Total	*4,021*	*87*	*47,616.1*	*251*
Significance	*0.080*	*0.056*		

C. Punitive-Compensatory Ratio

Alabama	39.04	1.96	127.3	17
Arkansas	2.48	1.50	2.7	8
California	16.87	8.78	27.0	12
Federal	22.89	1.64	154.6	86
Ohio	12.18	1.00	33.0	9
Texas	3.33	2.71	3.5	23
other states	65.61	1.52	424.0	96
Total	*37.21*	*1.81*	*279.5*	*251*
Significance	*0.439*	*0.399*		