

# Reconciling Punitive Damages Evidence

*Comment*

by

DANIEL E. HO\*

## *1 Introduction*

This comment on EISENBERG, HEISE, AND WELLS [2010] (henceforth EHW), explores how comparable data may have engendered sharply divergent views about punitive damages as (a) predictable or (b) erratic. Matched case control sampling provides an avenue to efficiently augment existing data with a key missing measure of the egregiousness of defendant conduct. I illustrate with a pilot study in Santa Clara County of the relationship between egregiousness and punitive damages.

## *2 The Puzzle of Scholarly Disagreement*

I am pleased to discuss EHW, a study of punitive damages that is admirable in many respects. The paper directly engages with the U.S. Supreme Court's interpretation of evidence published by EHW. Unlike much popular anecdotal coverage, it offers evidence based on a close to random sample of state court cases, and convincingly shows the limitations of the Court's analysis of raw mean and median statistics. Lastly, EHW applies modern statistical methods to demonstrate a robust positive correlation between compensatory and punitive damages (given positive punitives).

Exemplary in those respects, EHW and its conclusion that punitives are predictable based on the correlation also raise a puzzle over the broader, sharp empirical disagreement in the literature about punitive damages. One view posits that “[j]ury awards are highly *unpredictable*” (HERSCH AND VISCUSI [2004, p. 1]) and that the “dollar award often amounts to a stab in the dark” (SUNSTEIN et al. [2002, p. 31]). The contrary view is that punitives are *stable* and *predictable*, as EHW concludes that “the mass of punitive damages awards ha[s] been reasonably sober, modest in size, and without significant increases over time” (pp. 5f.). Such divergent conclusions stem in part from different evidentiary bases. See SUNSTEIN et al. [2002] (jury

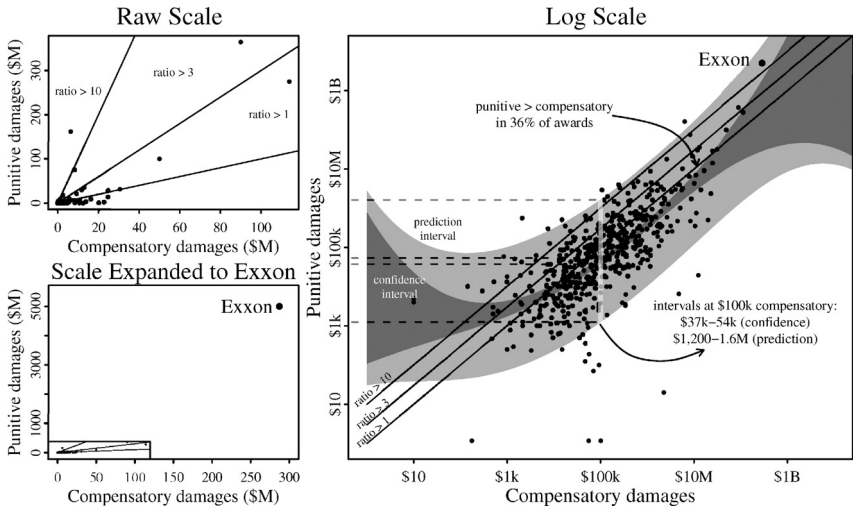
---

\* Stanford Law School. Thanks to Neal Ubriani for excellent research assistance, to Ted Eisenberg, Mike Heise, and Marty Wells for generously sharing data, and to John Donohue, Ted Eisenberg, Allie Dunworth, Jeff Fisher, and Stephen Galoob for helpful comments.

experiments) and EHW (observational surveys). Yet even based on the *same* data the disagreement appears to persist. The question is: why? I offer three conjectures to potentially reconcile these divergent views.

First, disagreement may stem from different quantities of interest. Theories diverge in normative criteria. Deterrence prioritizes the probability of detection, while jury standards emphasize egregiousness of conduct. Compensatory damages do not necessarily measure either. The correlation – assuming a punitive award, an *outcome* which itself may stem from rational or irrational decisions – may not speak to such quantities of interest (POLINSKY [1997], POLINSKY AND SHAVELL [1998]).

Figure 1  
Extrapolation, Transformations, and Uncertainty



Second, transformations may obscure substance. The upper left panel of Figure 1 plots the raw correlation between compensatory damages on the x-axis and punitives on the y-axis. The data are highly skewed, with significant outliers (hence, potential model dependence). Thin lines represent cutoffs for 1:1, 3:1, and 10:1 ratios. The bottom left panel expands the axes to include *Exxon Shipping Co. v. Baker*, 128 S.Ct. 2605 (2008) (not part of EHW’s data), far from the observed data, with an initial 17:1 ratio applying EHW criteria.<sup>1</sup> The right panel plots the same data on the log scale, with the same ratio lines. On the log scale, *Exxon* actually appears

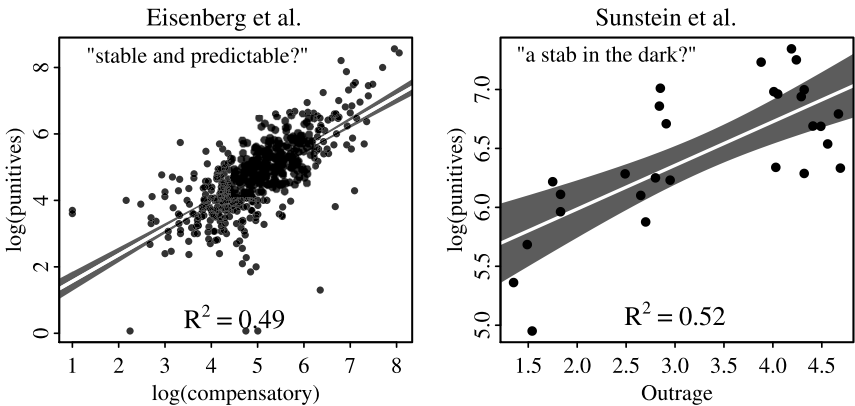
<sup>1</sup> There are potential limitations to these criteria, which rely on verdict-based punitive and compensatory damages. This, for example, ignores pretrial settlements and other payments made by Exxon. The district court of Alaska accordingly calculated a different ratio of 9.74:1. See *In re Exxon Valdez*, 296 F. Supp. 2d 1071 (D. Alaska 2004).

Delivered by Ingenta  
 ? 171.67.216.22 Fri, 01 Dec 2017 17:51:41  
 Copyright Mohr Siebeck

close to the trend – even small vertical deviations from the mass can signify large ratios. *Exxon*'s suggestion of a 1:1 "constitutional outer limit" calls into question some 36% of state court awards. Further, the dark grey band represents a 95% confidence interval (from a simple polynomial fit), comparable to one presented by EHW. Unpredictable punitive awards, however, may be less a matter of conditional expectation, but *variance*. The light grey band therefore plots the 95% prediction interval (which now contains *Exxon*). Substantively, the confidence interval for a \$100k compensatory award is \$37k–54k. The prediction interval, however, is considerably wider at \$1,200–1.6M.<sup>2</sup> Whether this variability is too high is a matter of interpretation.

Third, interpretation may be driven by priors. EHW, for example, argues that a high  $R^2$ , plotted in the left panel of Figure 2, demonstrates that punitives are predictable. The right panel plots the correlation from jury experiments between outrageousness of conduct and log of punitives (KAHNEMAN, SCHKADE, AND SUNSTEIN [1998]). That study concludes that while juries agree on outrageousness, the translation to dollar terms is just a "stab in the dark." Yet on the measure of  $R^2$ , the two are essentially identical. Whether a correlation is high is not a matter to be solved by fixed  $R^2$  thresholds, but by substantive knowledge.

Figure 2  
Divergent Interpretations Based on Comparable  $R^2$



### 3 A Way Forward? Matched Case Control Sampling

While EHW's body of work sheds considerable insight into the tort system, major questions remain. Setting aside deterrence rationales, are juries consistent in translating egregiousness of conduct into punitive damages? While EHW's data

<sup>2</sup> This point is recognized in EISENBERG et al. [1997, p. 657].

Delivered by Ingenta  
 ? 171.67.216.22 Fri, 01 Dec 2017 17:51:41  
 Copyright Mohr Siebeck

contains a host of covariates, it lacks key measures of such case facts. One way forward consists of *case control sampling*. By matching punitive cases with all cases comparable on all dimensions but egregiousness, we may efficiently study the relationship in actual cases. To illustrate this potential, we retrieved case files for punitive cases in Santa Clara from EHW's data matched with all other cases on bench/jury trial and case type. (The process was disturbingly time-consuming, involving, *inter alia*, a bailiff apprehending a research assistant.) Due to sampling and incomplete matching, this resulted in 14 cases summarized in Table 1. Non-italicized cases have positive punitive damages while italics indicate matched cases. Other columns present covariates (trial type, case type, year, compensatory award size, whether punitive damages were sought, and punitive award size). Four tentative lessons emerged.

Table 1  
Summary of Matched Case Control Sample

Pun. Cases <i>Matched Cases</i>	Trial type	Case type	Year	Comp. (\$)	Pun. sought	Pun. (\$)
Larson v. Dutra	jury	int. tort	2005	56,500	yes	20,000
<i>Landin v. Cloud Nine</i>	jury	int. tort	1996	58,581	yes	2,500
<i>Baumell v. Demasi</i>	jury	int. tort	1996	36,000	no	
<i>Movahedi v. Bank of West</i>	jury	int. tort	1996	75,000	yes	0
<i>Crowder v. Branam</i>	jury	int. tort	1996	750,000	yes	0
Warden v. Moore	jury	fraud	2005	617,417	yes	308,667
<i>Carlbom v. Drs. Billing</i>	jury	fraud	2005	267,085	no	
Baron v. Fire Insurance	jury	contract	2005	156,462	yes	1,500,000
<i>Dalton v. Century 21</i>	jury	contract	2005	85,495	yes	0
<i>Katz v. Stewart</i>	jury	contract	2001	31,000	yes	0
Rehon & Roberts v. Mahl	bench	fraud	2001	749,572	yes	150,000
<i>ACS v. Max Group</i>	bench	fraud	2001	260,798	yes	0
<i>Peterson v. Perez</i>	bench	fraud	2001	96,500	yes	0
<i>Lee v. Van Blitter</i>	bench	fraud	2001	52,331	yes	0

(1) *Case Incomparability*. Clerks are instructed to fill out “case types,” but these are measured with significant error. One case (*Baumell*), for example, involved a breach-of-contract claim by a jewelry supplier, yet the case was classified as an intentional tort. Claim types in EHW are only one of all causes of action in a complaint and do not necessarily represent the claims on which a plaintiff prevailed.

(2) *Lawyer Gatekeeping*. Inferences from the EHW's data are not necessarily about jury behavior, as often interpreted. One case (*Warden*) involved an attempt to secure

a roughly \$600k loan from parents by manufacturing a fake grandchild, resulting in over \$300k in punitives. The matched fraud case (*Carlblom*) involved the failure by a doctors billing service to properly process claims. Punitives were simply not sought. If lawyers, as repeat players, serve a gatekeeping role, they may impose consistency on the system even where juries might otherwise be erratic.

(3) *Egregiousness*. Matched cases reveal clear instances where egregiousness distinguishes a punitive damage case. In one case (*Larson*), the plaintiff was waiting in her car and the defendant “opened the door to the passenger side and ... repeatedly beat [her] about the head with a black metal flashlight approximately twenty times,”<sup>3</sup> resulting in \$56k in compensatory and \$20k in punitive damages. In a matched case (*Landin*), with \$55k in compensatory damages but only \$2.5k in punitives, behavior was by plausible accounts less egregious (although far from commendable). A limo driver “kicked [customer plaintiff] in the testicles ... and [his] leg out from under him. Plaintiff landed on his foot in an awkward manner and fell to the ground.”<sup>4</sup>

(4) *Observational Equivalence*. California special verdicts often permit reconstruction of specific facts found by the jury. Cal.C.C.P. § 625. But the record remains thin. Many cases are consistent both with wildly erratic and eminently rational jury behavior. Cold records pose comparable challenges both to academic research and appellate review.

I conclude with two thoughts. First, one of the major benefits of serious empirical inquiry into the tort system is that it imposes conceptual clarification. Case control sampling, for example, challenges conceptions of “egregiousness” when the aim is to identify cases identical in all exogenous respects other than egregiousness.

Second, while the case law prior to *Exxon* eschewed any “simple mathematical formula” and used ratios only as a “raised eyebrow” test,<sup>5</sup> *Exxon*’s empirical turn must come with caution. EHW shows that a 1:1 ratio may affect a wide range of cases. At the same time, “predictability” may depend critically on prior beliefs about the jury system given cold records. Case-by-case appellate inquiry may be more warranted when the function of judicial review may effectively be one of outlier detection. And with burgeoning studies, such as EHW, the empirical knowledge base and consensus is just starting to grow.

### References

EISENBERG, T., J. GOERDT, B. OSTROM, D. ROTTMAN, M. T. WELLS [1997], “The Predictability of Punitive Damages,” *The Journal of Legal Studies*, 26, 623–661.

<sup>3</sup> Declaration of David Kraft at 3, *Larson v. Dutra*, No. 103-CV007400 (Super. Ct. Santa Clara Jan. 31, 2005).

<sup>4</sup> Plaintiff’s Trial Brief at 3, *Landin v. Cloud Nine Limousine Service*, No. CV 735263 (Super. Ct. Santa Clara May 30, 1995).

<sup>5</sup> *BMW of North America, Inc. v. Gore*, 517 U.S. 559, 582, 613 (1996).

- , M. HEISE, AND M. T. WELLS [2010], “Variability in Punitive Damages: Empirically Assessing *Exxon Shipping Co. v. Baker*,” *Journal of Institutional and Theoretical Economics*, 166, 5–26.
- HERSCH, J., AND W. K. VISCUSI [2004], “Punitive Damages: How Judges and Juries Perform,” *The Journal of Legal Studies*, 33, 1–36.
- KAHNEMAN, D., D. SCHKADE, AND C. R. SUNSTEIN [1998], “Shared Outrage and Erratic Awards: The Psychology of Punitive Damage Awards,” *Journal of Risk and Uncertainty*, 16, 49–86.
- POLINSKY, A. M. [1997], “Are Punitive Damages Really Insignificant, Predictable, and Rational? A Comment on Eisenberg et al.,” *The Journal of Legal Studies*, 26, 663–677.
- AND S. SHAVELL [1998], “Punitive Damages: An Economic Analysis,” *Harvard Law Review*, 111, 869–962.
- SUNSTEIN, C. R., R. HASTIE, J. W. PAYNE, D. A. SCHKADE, AND W. K. VISCUSI [2002], *Punitive Damages*, The University of Chicago Press: Chicago.

Daniel E. Ho  
Stanford Law School  
Stanford University  
559 Nathan Abbott Way  
Stanford, CA 94305  
U.S.A.  
E-mail:  
dho@law.stanford.edu