JIPEL

NYU Journal of Intellectual Property
& Entertainment Law

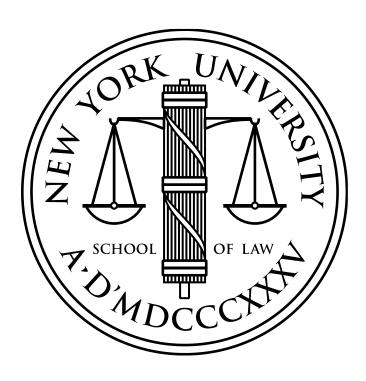## Statement of Purpose

Consistent with its unique development, The New York University Journal of Intellectual Property & Entertainment Law (JIPEL) is a nonpartisan periodical specializing in the analysis of timely and cutting-edge topics in the world of intellectual property and entertainment law. As NYU's first online-only journal, JIPEL also provides an opportunity for discourse through comments from all of its readers. There are no subscriptions, or subscription fees; in keeping with the open-access and free discourse goals of the students responsible for JIPEL's existence, the content is available for free to anyone interested in intellectual property and entertainment law.

A nonpartisan periodical, the Journal is committed to presenting diverse views on intellectual property and entertainment law. Accordingly, the opinions and affiliations of the authors presented herein do not necessarily reflect those of the Journal members.

The Journal is also available on WESTLAW, LEXIS-NEXIS and HeinOnline.

DATA VALIDATION PROCEDURES USED FOR BEEBE ARTICLE

VOLUME 10 EDITORIAL BOARD OF THE NYU JOURNAL OF INTELLECTUAL PROPERTY
AND ENTERTAINMENT LAW (JIPEL)

In order to give academics and practitioners greater assurance in the results found in Professor Beebe's article, JIPEL assigned several staff editors to review the author's data coding across a representative sample set of cases for a population of this size. This set was determined to be 35 cases through consultation with a publicly available sample size calculator.[1]

Several editors were then assigned to verify data coding for the following variables: disposition, reversal and appeal, venue, the outcome of each fair use factor, whether subfactor considerations were cited, the word count devoted to each fair use factor and fair use overall, and the number of times an opinion cited to a specific court. A senior editor then verified these results and tabulated the below output table.

JIPEL focused its data validation on data coding for variables that the author either discussed in detail in his article or which were associated with significant results. JIPEL did not review the data coding for third-party data sources, such as judge characteristics including race, gender, ideology, and partisan leanings. Professor Beebe provides the underlying sources for this data in the relevant portion of his analysis. JIPEL did not validate data coding for variables not utilized in the article.

Through reviewing this sample set of cases, JIPEL was able to verify that the coding of the overall population of cases analyzed by the author had a margin of error of less than 15 percentage points at a 95% confidence level. JIPEL calculated the p-values associated with 21 categories of data and 131 sub-categories to substantiate that there was no reason to reject the assumption that the population effect was accurate.[2] Overall, the staff reviewed close to 1500 data inputs and found very few errors.

---

[1] *Sample Size Calculator*, CLINCALC, https://clincalc.com/stats/samplesize.aspx (lasted visited Nov. 12, 2020).

[2] P-values were specifically checked for case dispositions, reversal and appeal rates, treatment of each of the four fair use factors, and treatment of most sub-factors. A p-value calculates the likelihood that a random sample of the same size as the current sample would have a difference between the population effect and the sample effect that is equal to or greater than the

Finally, JIPEL is providing the underlying data and data key used by Professor Beebe in his analysis. The professor encourages feedback and collaboration and has agreed to this data sharing full-heartedly.

    (1) Professor Beebe's underlying data coding (excel)
    (2) Professor Beebe's data key (pdf)

We acknowledge that this method is not suitable for every empirically-focused article. Nevertheless, we believe that making it available may help other journals move forward along the path toward adopting more rigorous and standardized review for the underlying data and assumptions in empirical legal works.[3]

TABLE 1: JIPEL DATA VALIDATION FOR REPRESENTATIVE SAMPLE SET OF 35 CASES

Variable	Total Entries Recorded by Author	Mistakes Found by JIPEL	Calculated Percent Error	P-value <0.05 for any sub-category
Disposition of the case	35	1	2.86%	1 of 17
Reversal and appeal	35	0	0%	0 of 5
Outcome of each fair use factor	140	3	2.14%	0 of 27
Venue	35	1	2.86%	N/a
Whether subfactors considerations were cited in the analysis of each fair use	560	9	1.61%	7 of 82[4]
Word count devoted to each fair use factor and fair use analysis as a whole	175	1	0.57%	N/a
Number of times an opinion cited to a specific Circuit Court or the SDNY	490	7	1.43%	N/a

current measured difference under the assumption that the population effect is accurate. Of the 131 p-values reviewed, only 8 p-values were <0.05, suggesting that there is no reason to reject the assumption that the population effect was accurate. *Introduction to Power Analysis*, UCLA https://stats.idre.ucla.edu/other/mult-pkg/seminars/intro-power/ (Last visited Nov. 12, 2020).

   [3] A special thanks to a friend of JIPEL, economist Alissa Dubnicki Ph.D., for her assistance and advice in helping JIPEL to architect this data validation exercise.

   [4] JIPEL only checked p-values for sub-factors relied on in analysis, although data validation checked all categories. P-values checked included 15 sub-factors across four fair use factors and "other," a catch-all to consider whether factors besides the four factors was used, as well as bad faith.