Appendix to "Politics and Rulemaking at the Copyright Office"

Gabriel J. Michael

Data File

The dataset used in the paper is available for download as a standard CSV file. This file can be imported into a spreadsheet or statistical analysis program. The file contains the following fields:

- *cycle* (numeric): specifies the year of the rulemaking: 2000, 2003, 2006, 2010, or 2012.
- *comment_type* (string): specifies whether a comment was an initial or a reply comment. In this dataset, all comments are initial comments.
- *doc_num* (numeric): an internal reference number for the comment. In early cycles, this corresponds with naming conventions on the Copyright Office website.
- *name* (string): the name or names of the individuals or organizations submitting the comment.
- *type* (category): a coded variable indicating the commenter type. One of business, education, individual, or NGO.
- *subtype* (category): an unused coded variable indicator the commenter subtype.
- *represented* (binary): a coded variable indicating whether the commenter had legal representation.
- *valid* (binary): a coded variable indicating whether the comment proposed a specific class of copyrighted works to be exempted.
- *exempt* (string): a list of exemptions requested in the comment.
- *renew* (binary): a coded variable indicating whether the comment requested an exemption that had been granted in an immediately prior cycle. Coded as "na" for the 2000 cycle, since there were no prior cycles.
- *granted* (binary): a coded variable indicating whether any of the exemptions requested in a comment were granted by the Copyright Office.
- *pages* (numeric): specifies the number of pages in a comment.
- *city* (string): when available, indicates the city of the commenter.
- *state* (string): when available, indicates the state of the commenter.
- *country* (string): when available, indicates the country of the commenter.
- *note* (string): various notes.

Data Sources

The full text of every comment submitted during each rulemaking is available online at the Copyright Office website. Direct links to pages for the individual cycles can be found below:

- 2012: http://www.copyright.gov/1201/
- 2010: http://www.copyright.gov/1201/2010/
- 2006: http://www.copyright.gov/1201/2006/
- 2003: http://www.copyright.gov/1201/2003/
- 2000: http://www.copyright.gov/1201/anticirc.html

During data collection, I downloaded data from these pages, which I have organized into a complete collection of both initial and reply comments for all five cycles. This reproduction is for research purposes, and thus is a fair use. Due to the fact that copyright law protects comments submitted during rulemakings, I do not provide public access to this collection (some of the organizations submitting comments are highly protective of their intellectual property, and might object to their comments being republished). If you wish to use the collection for research purposes, please contact me directly at gmichael@gwu.edu.

The pages linked above also contain the Federal Register notices and recommendations describing which exemptions were granted during each cycle.

Coding

As noted in the paper, I did all the coding myself, so there are no issues of inter-coder reliability.

All "types" other than individual imply that the comment originated from some sort of corporate body. When comments were jointly submitted, I code based on the first or most prominently listed commenter. If an individual mentions that they hold an executive level position at a business, I code them as a business, unless they note that their views are their own.

Commenters that are lawyers, or hold positions like legal director, general counsel, vice president for legal affairs, etc. are coded as having representation. This is also true for comments originated from law firms or law clinics. In cases where a comment did not explicitly identify the writer as a lawyer, but there seemed to be a high likelihood of this being the case, I did research on the commenter to assist in coding the variable. Coding of this variable could result in false negatives since not all comments drafted or aided by lawyers readily identify those lawyers. However, this would bias the variable away from significance in the hypothesized direction.

Comments filed jointly are considered as a single comment. This implies that if any member of the joint commenters had legal representation, the entire comment is considered to have legal representation.

Where two cities are specified, I use the first, except when the first is a law office and the second is an organizational headquarters. In that case, I use the second.

I originally coded whether an exemption had been granted as an ordinal variable: for comments requesting multiple exemptions, the value indicated the number of requested exemptions that were granted. Only a few comments had values greater than 1, so the ordinal variable was collapsed into a binary variable.

In the 2012 cycle, the Copyright Office received only five individual submissions from commenters without legal representation. In previous cycles, these comments would have been coded as invalid. However, in this cycle, it appears that the Copyright Office read these submissions (all of which are clearly written by non-professionals) and amalgamated their concerns into a single, broad proposed exemption.

Correspondence Between Requested and Granted Exemptions

In the paper, I note that establishing a link between exemptions requested by commenters and exemptions granted by the Copyright Office is trivial due to the nature of the rulemaking: the office does not propose its own exemptions, but instead relies solely on proposals originating from commenters.

To illustrate the correspondence between requested and granted exemptions, I have selected five illustrative examples, one from each rulemaking cycle. While the level of textual correspondence between the requests and the exemptions varies from similarity to near identity, there is always a clear connection between the request and the granted exemption.

- 2000

Spectrum Software submitted a comment containing the following information:

> "If a company goes out of business, there is no one to support the authorized customer when a hardware lock is damaged since it is the manufacturer of the software, not the manufacturer of the hardware lock, that is the only one that can program the dongle… In another example, once a company has been acquired, their software program is phased out. After a period of time, the program and lock device is no longer supported because companies want the customer to upgrade to the newer combined product or they are using a different hardware lock device."

The following exemption was granted:

> "Literary Works, Including Computer Programs and Databases, Protected by Access Control Mechanisms That Fail to Permit Access Because of Malfunction, Damage or Obsoleteness" Clarifying text states "Substantial evidence was presented on this issue, in particular relating to the use of ''dongles,'' hardware locks attached to a computer that interact with software programs to prevent unauthorized access to that software."

- 2003

The American Foundation for the Blind submitted a comment containing the following information:

> "In these comments we address the class of all literary digital publications, such as e-books…. It certainly cannot preserve the rights of people who are blind or visually impaired to use their synthetic speech and braille devices to access secured content… We urge the Librarian to reaffirm the long history of copyright law and rulings which explicitly affirm the right of persons who are blind or visually impaired, as well as those providing reading matter to them, to repurpose content into accessible formats, including such content as may be secured by digital means."

The following exemption was granted:

> "Literary works distributed in ebook format when all existing ebook editions of the work (including digital text editions made available by authorized entities) contain access controls that prevent the enabling of the ebook's read-aloud

function and that prevent the enabling of screen readers to render the text into a ``specialized format.''"

- 2006

The Wireless Alliance requested the following exemption:

"Computer programs that operate wireless telecommunications handsets. (Mobile firmware)… Phone unlocking software is a tool that can circumvent the software locks carriers use to stop customers from using the handsets they purchase on competing mobile networks. Using a mobile handset on a different network is clearly non-infringing activity."

The following exemption was granted:

"Computer programs in the form of firmware that enable wireless telephone handsets to connect to a wireless telephone communication network, when circumvention is accomplished for the sole purpose of lawfully connecting to a wireless telephone communication network."

- 2010

Professor Renee Hobbs of Temple University requested the following exemption:

"Audiovisual works that illustrate and/or relate to contemporary social issues used for the purpose of teaching the process of accessing, analyzing, evaluating, and communicating messages in different forms of media."

The following exemption was granted:

"Motion pictures on DVDs that are lawfully made and acquired and that are protected by the Content Scrambling System when circumvention is accomplished solely in order to accomplish the incorporation of short portions of motion pictures into new works for the purpose of criticism or comment, and where the person engaging in circumvention believes and has reasonable grounds for believing that circumvention is necessary to fulfill the purpose of the use in the following instances: … Educational uses by college and university professors and by college and university film and media studies students."

- 2012

The Electronic Frontier Foundation requested the following exemption:

"Computer programs that enable wireless telephone handsets ("smartphones") and tablets to execute lawfully obtained software applications, where circumvention is undertaken for the purpose of enabling interoperability of such applications with computer programs on the handset or tablet."

The following exemption was granted:

"Computer programs that enable wireless telephone handsets to execute lawfully obtained software applications, where circumvention is accomplished for the sole purpose of enabling interoperability of such applications with computer programs on the telephone handset."

R Code for Tables and Figures

The R code used to generate the tables and figures in this paper is available for download as a standard text file with a ".R" extension. It has Unix-style newlines, so you may need to change settings to get it to display properly in some Windows programs. It requires the following packages and their dependencies: xtable, pscl, texreg, and ggplot2. The code was developed on R version 2.15.2 running on x86_64-apple-darwin9.8.0/x86_64 (64-bit).

The post-estimation plots use a file of test data that contains 1600 rows, containing combinations of variables. This file is also available for download as a standard CSV file.

Please note that this code will not run as is. In addition to a working R environment, you need to have both the data and test data files, specify the path to these files, and uncomment the "read.csv" lines.

Predictions Based on Model 4 (2000-2010 Data Only)

As discussed in the paper, I used a model based on 2000 to 2010 data to generate predictions for the 2012 cycle in advance of the Copyright Office's final decisions. In order to timestamp my predictions and prove they were made prior to the announcement of exemptions, I tweeted them from my Twitter account.

Simply timestamping predictions using Twitter is insufficient to serve as evidence of prediction, since users can easily create multiple Twitter accounts, issue different predictions via each account, and then *post facto* decide which account and predictions to cite. To address this issue, I made sure to use my named, public Twitter account (@gabrieljmichael). This account is clearly connected to me, meaning I could not later disclaim the predictions.

Each prediction specified the year and number of the exemption, whether it was likely to be accepted/renewed or rejected, and included a link to the full text of the exemption.

The text of the predictions (timestamped June 30, 2012) were as follows:

2012 proposed #dmca #1201 exemptions 2, 5, and some form of 6[ABC] likely to be accepted or renewed #prediction http://goo.gl/dPk6x

https://twitter.com/gabrieljmichael/status/219159834296201217

2012 proposed #dmca #1201 exemptions 1, 10B likely to be rejected #prediction http://goo.gl/dPk6x

https://twitter.com/gabrieljmichael/status/219160033064259584

2010 #dmca #1201 exemptions 4 (games), 5 (dongles) likely to be discontinued #prediction http://goo.gl/eXkfO

https://twitter.com/gabrieljmichael/status/219220970395799552

To summarize, I predicted that during the 2012 rulemaking cycle:

- Exemption 2 of 2012 (requesting exemption of literary works containing DRM that prevents access by blind persons) would be accepted.

- Exemption 5 of 2012 (requesting exemption of software used for "jailbreaking" or "rooting" of mobile phones or tablets) would be accepted.
- Some form of Exemption 6A, 6B, or 6C (similar proposals requesting exemption of software used to "unlock" mobile phones locked to a single network provider) would be accepted
- Exemption 1 of 2012 (requesting exemption of public domain literary works) would be rejected.
- Exemption 10B of 2012 (requesting a broad exemption for backup, format shifting, and other purposes) would be rejected.
- Exemption 4 of 2010 (an exemption granted in 2010 permitting security research on video games) would not be renewed.
- Exemption 5 of 2010 (an exemption granted in 2010 exempting computer programs protected by hardware dongles) would not be renewed.

These predictions were obtained by using the coded data of initial comments in the 2012 cycle. The cutoff point for my predictions were commenters with z-scores above 1.16, corresponding to about an 88% probability of an exemption being granted.

The model predicted that the following commenters had a high likelihood of receiving an exemption:

- American Council for the Blind and American Foundation for the Blind (filing jointly)
- Electronic Frontier Foundation
- Software Freedom Law Center
- Consumers Union
- International Documentary Association, Kartemquin Educational Films, Inc., National Alliance for Media Arts and Culture, and Independent Filmmaker Project (filing jointly)
- Telecommunications for the Deaf and Hard of Hearing, Inc., Gallaudet University, and Participatory Culture Foundation (filing jointly)
- Public Knowledge

These commenters had requested 13 unique exemptions. However, because the model can only predict commenter influence, and some comments requested multiple exemptions, I further limited my predictions to those noted above.

When the Copyright Office announced its exemptions on October 26, 2012, its decisions corresponded to all of the predictions discussed above except one: rather than unqualifiedly renewing some form of exemption 6A/6B/6C, it instead renewed the phone unlocking exemption for only 90 days, effectively rejecting the exemption.