

Copyright Renewal of U.S. Books Published in 1932: Reanalyzing Ringer’s Study to Determine a More Accurate Renewal Rate for Books

Jamie Carlstone, Ayla Stein, Michael Norman, and John Wilkin

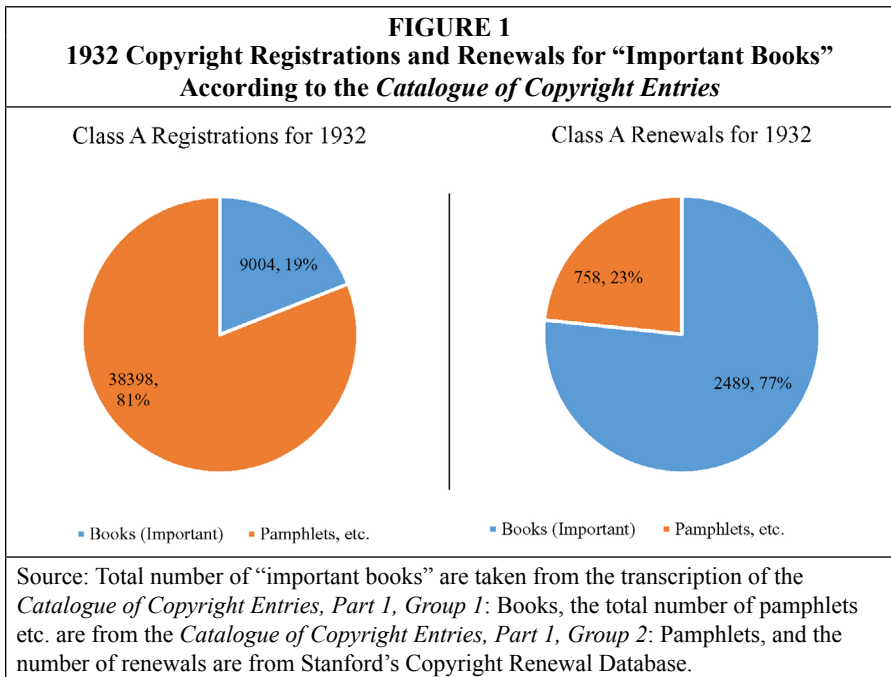
In 1961, Barbara Ringer published “Study No. 31: Renewal of Copyright,” where she determined the renewal rate for fiscal year 1932 U.S. publications. In that study, she concludes that the renewal rate for Class A works for FY1932 was 7 percent. This paper seeks to reassess Ringer’s study by analyzing the copyright registrations for 1932 and their renewals published in the *Catalogue of Copyright Entries*. This was done to determine a renewal rate specifically for books rather than Class A as a whole, which includes other materials. The analysis determines that the copyright renewal rate for books is actually 26 to 33 percent, significantly higher than 7 percent claimed by Ringer.

Introduction

In 1960, Barbara Ringer of the United States Copyright Office published “Study No. 31: Renewal of Copyright,” to “analyze the subject of renewal copyright in all its aspects—duration, formality, and ownership—in order to form the basis for an objective determination of renewals’ value or lack of it.”¹ Ringer’s report determined that fiscal year 1932 Class A works were renewed 7 percent of the time. Class A, as defined by the Copyright Act of 1909, is described as “Books including composite and cyclopaedic works, directories, gazetteers, and other compilations.”² However, as there are many other types of works included in Class A for 1932 (including pamphlets, advertisements, and newspaper articles), it is not possible to determine the renewal rate for books by just looking at Class A totals, as Ringer did, because books are likely to be renewed more often than pamphlets and other materials. In fact, approximately 19 percent of the registrations in the *Catalogue of Copyright Entries* for Class A were called “important books,” but the “important books” account for 77 percent of the renewals (see figure 1).³

*Jamie Carlstone is E-Serials Bibliographic Control Specialist, Ayla Stein is Metadata Librarian, Michael Norman is Head of Content Access Management (all in the Department of Content Access Management of the Main University Library), and John Wilkin is the Juanita J. and Robert E. Simpson Dean of Libraries and University Librarian, all at University of Illinois at Urbana-Champaign; e-mail: jdc6@illinois.edu, astein@illinois.edu, manorman@illinois.edu, jpwilkin@illinois.edu. ©2018 Jamie Carlstone, Ayla Stein, Michael Norman, and John Wilkin, Attribution-NonCommercial (<http://creativecommons.org/licenses/by-nc/4.0/>) CC BY-NC.

Ringer’s assessment of Class A has led to 7 percent often being cited as the renewal rate for books for this period. As John Wilkin demonstrated in “How Large Is the ‘Public Domain’?: A Comparative Analysis of Ringer’s 1961 Copyright Renewal Study and HathiTrust CRMS Data,” while the registration rate for Class A FY1932 works is 7 percent, we cannot then conclude that the renewal rate for books is this low.⁴ A higher renewal rate supports the conclusion of HathiTrust’s Copyright Review Management System (CRMS), which cites 51 percent as the renewal rate for HathiTrust Digital Library Materials.⁵ An accurate renewal rate is critical for organizations participating in digitization projects because underestimating the number of works in the public domain can make them vulnerable to legal risk by inadvertently opening up copyrighted works. Knowing the size of the public domain is necessary for digitization project planning so that libraries and other organizations can responsibly make copyright status determinations. It is also important for future conversations about revisions to copyright law and the importance of renewal. The copyright renewal rate for books published in the United States in 1932 is more likely between 26 and 33 percent, a number that stands in stark contrast to the 7 percent rate often cited for books for 1923–1963.



The Copyright Office used the term “book” both narrowly (how an American consumer would likely define a book, such as published monographs and novels of a certain length that have market value) and broadly (everything copyrighted as Class A, including other formats like telephone directories and pamphlets). Class A, as discussed in this paper, will strictly refer to the Copyright Act of 1909 definition, which includes the whole class encompassing a broad range of materials. This is an important distinction to make; for example: *Brave New World* by Aldous Huxley⁶ has the same copyright class as an advertisement for tobacco called, “Which? 15¢ and no smoke? etc.” by the Brown and Williamson Tobacco Corp.⁷ Any statistics kept by the Copyright Office for Class A categorizes these two fundamentally different types of works as the same, so it is important to take this into account.

Moreover, Ringer's study only collected data for FY1932 works, but the renewal rate of books likely increased steadily after that year. She recognized this limitation as well; according to Ringer, "In 1910 renewal was made for about 3 1/2% of the works copyrighted in 1883; this percentage had doubled by 1940 and has now redoubled. Last year the total of copyrighted works renewed was nearly 15%."⁸ Therefore, we cannot attribute a low renewal rate in 1932 to subsequent years.

The year 1932 is still a significant date for copyright purposes because it is between 1923 and 1963, thus reflecting a period of time where the public domain status of works is complicated by changes in copyright law. During this period, a work could fall into the public domain if the copyright owner failed to put a copyright notice on the work, failed to register the work with the Copyright Office within 5 years after publication, or failed to properly renew the work with the Copyright Office after the first 28-year term. As Peter Hirtle explains, the Sonny Bono Copyright Act in 1998 extended the length of copyright to 95 years for works published before January 1, 1978, if the materials were renewed when they originally came up for renewal pursuant to the law at that time. Since copyright could be renewed after 28 years, any registrations made in 1932 had to be renewed 28 years later (or in 1960) to remain under copyright protection today. The extension means that materials published in 1923 could be protected under copyright for 95 years, or until 2019, if they were renewed. Otherwise, the material would currently be in the public domain with a copyright expiration date of 1998 (the date of the Sonny Bono Act). After 2019, the date will start to shift each year (for example, works published in 1924 are protected until 2020). Renewals for 1932 works will be protected by copyright until 2027.⁹ This is why finding a renewal rate for this time period is significant—many early-to-mid 20th century works, if they were renewed, are not currently in the public domain.

In this follow-up study to Wilkin, we examine the copyright registrations from 1932 and the renewals for those registrations 28 years later to establish a more accurate renewal rate for books published in the United States in 1932. To do this, we also analyze Copyright Office documentation and sources of copyright data to assess the challenges in creating a methodology for calculating a more accurate renewal rate. This is made complicated by the main source for copyright registrations/renewals, the *Catalogue of Copyright Entries (CCE)*. The CCE is an invaluable, yet flawed, bibliographic resource. Understanding the CCE is so important to understanding copyright and renewal that a detailed analysis of the CCE is included in this report. This study also looks at data from the Copyright Office's *Annual Report* for 1932 and 1933, as well as the Stanford Renewal Database, which contains transcriptions of the CCE renewals. Having a clearer sense of the renewal rate for 1932 books can help us better understand the likely size of the public domain for books published in the United States between 1923 and 1963.

The Copyright Office and "Books"

The Copyright Office uses several categorical terms to define books: "important books" and "books proper." "Important books" is a description used for the works entered into the *Catalogue of Copyright Entries* volume, *Part 1 Group 1: Books*. "Books proper" is the count of books listed in a table in the Copyright Office *Annual Report*.¹⁰ There are more "books proper" than "important books," underlining the fact that these terms should not be equated. However, overlap between the two is certain (for instance, something that is an "important book" could also be a "book proper"). Both are valuable data points that will be used for renewal rate calculations.

The Copyright Office and the *Catalogue of Copyright Entries*

The recording of copyright registrations is the responsibility of the United States Copy-

right Office, established in 1897 as a part of the Library of Congress.¹¹ The Copyright Office published entries of copyright registrations in a serial, *The Catalogue of Copyright Entries*. The CCE began publication in 1891 under the earlier title *The Catalogue of Title-Entries*.¹² The CCE was produced by the Copyright Office but was originally paid for by the Department of the Treasury because the CCE was conceived as a tool to aid customs officers in the prevention of copyright infringement. In 1904, the Secretary of the Treasury recommended the financial responsibility for the CCE shift to the Library of Congress, since it was not actually useful to customs officers. Thorvald Solberg, then head of the Copyright Office, believed in its value as a bibliographic tool and record of the United States book trade and so the publication was continued and fiscal responsibility for it transferred to the Library of Congress on June 30, 1906.¹³ From this point on, responsibility for the CCE was entirely within the purview of the Library of Congress. Over the decades that the CCE was published, both its organization and the format of entries changed. Copyright registrations (and by extension, the Copyright Office and CCE) were also affected by fee increases in 1928 and 1948, as well as global events such as the Great Depression and World War II.¹⁴ This paper will focus on data from 1932, unless otherwise noted; therefore, some of our analysis will not apply to the practices of the CCE during its entire lifespan.

In 1932, the four parts of the CCE, including renewals and indexes, were published in six volumes. It should be noted that the registration classes, A-M and R (renewals) (see table 1),¹⁵ do not correspond to the structure of the CCE publication. Instead of publishing each class in its own volume, the Copyright Office consolidated the 23 classes in 4 parts (with part 1 split into three groups):

- Part 1, Group 1: Books (*P1G1: Books*)
- Part 1, Group 2: Pamphlets, contributions to Newspapers or Periodicals, Lectures, and Maps (*P1G2: Pamphlets*)
- Part 1, Group 3: Dramatic Compositions and Motion Pictures (*P1G3: Dramatic Compositions*)
- Part 2: Periodicals
- Part 3: Musical Compositions (*P3: Musical Compositions*)
- Part 4: Works of Art, Photographs, Prints and Pictorial Illustrations¹⁶

Class A materials were published across multiple volumes of the CCE; these are the volumes relevant to our discussion. Most are in Part 1, Groups 1–2, but a few Class A entries appear in *Part 1, Group 3: Dramatic Compositions*¹⁷ and *Part 3: Musical Compositions*.¹⁸ Occasionally, identical Class A entries will appear in more than one place, such as

TABLE 1
Copyright Registration Classes

	Subject Matter of Class
A	Books
	Books proper
	Pamphlets, Leaflets, etc
	Contributions to newspapers and periodicals
B	Periodicals
C	Lectures, sermons, addresses
D	Dramatic or dramatic-musical compositions
E	Musical compositions
F	Maps
G	Works of art, models or designs
H	Reproductions of works of art
I	Drawings or plastic works of a scientific or technical character
J	Photographs
K	Prints and pictorial illustrations
L	Motion-picture photoplays
M	Motion pictures not photoplays
R	Renewals

Source: Copyright Office's *Annual Report* for 1932

the dramatic composition *Plays & Poems* by W.S. Gilbert, which appears in both *PIG1: Books*¹⁹ and *PIG3: Dramatic Compositions*.²⁰ Additionally, the Copyright Office reused registration IDs for unrelated titles.²¹ Therefore, registration IDs and entries cannot be assumed to be unique.

For the purposes of this research, we are focused on American copyrights; this means that we have intentionally excluded foreign and ad interim²² copyright registrations from the discussion (those with AF and Ad Interim prefixes) and will not include them when discussing the number of entries or renewals.

Part 1, Group 1: Books

For the purpose of this analysis, *Part 1, Group 1: Books (PIG1: Books)* for 1931–1933 was transcribed into a spreadsheet. The transcription was contracted out to a vendor.²³ For each entry, the author, author birth/death dates, relationship designator, edition, title, publication date, publisher, publication year, copyright date, receipt date, claimant, registration ID, and Library of Congress number were transcribed, if present. After the transcription was received, the part of the *CCE*, fascicle number, and fascicle spanning year were added by the authors of this paper. If an entry had multiple registration IDs, only the first was transcribed.

Type	Table, Page ii	Transcription
(a) United States publications	8,994	9,004
(b) Foreign books in foreign languages	2,471	2,488
(c) Foreign books in English language under Ad Interim	1,452	1,408
Total	12,917	12,900
Source: <i>Catalogue of Copyright Entries, Part 1, Group 1: Books</i> for 1932 and its transcription.		

Each copyright registration in *PIG1: Books* contains bibliographic information, a registration ID, and a catalogue entry number. The *PIG1: Books* issues were published every other day “In order to secure to librarians the promptest possible information as to the American and foreign books deposited for copyright.”²⁴ The 1932 *PIG1: Books* is made up of 152 fascicles printed every other day between February and December of 1932, along with an annual index. It would be up to libraries to bind the fascicles with the renewals and indexes, so many physical copies of the *CCE* are bound differently. The entries are alphabetical by author surname but start over in each fascicle; without knowing the date of an entry, the index is necessary for locating a specific entry. The entry number is continuous throughout the fascicles and is used in the index for locating entries. *PIG1: Books* includes a summary table of the total number of registrations in the volume on page ii. The total number of entries of Class A materials in the summary table in *PIG1: Books* 1932 is 8,994;²⁵ however, we find 9,004 entries according to our transcription (see table 2).²⁶ For consistency, this paper will use the transcription numbers for the total. Figure 2 and table 3 show how the *CCE* entries were transcribed into a table. In the example, the main heading (Aldous Huxley) is the author; a title is used in entries where no author is available. The registration ID

in the example is A-47485. The “A” designates its class. The entry number is 759, so it is the 759th entry in the volume. The book has also been given a Library of Congress Classification Number (LCCN): 32-3525.²⁷

FIGURE 2
Example of a 1932 PIG1: Books Entry, Brave New World by Aldous Huxley

Huxley, Aldous Leonard, 1894–
 Brave new world, a novel by Aldous Huxley. Garden City, N. Y., Doubleday, Doran & company, inc., 1932.
 5 p. l., 311 p. 22½cm.
 “First edition.”
 “This edition is limited to two hundred and fifty copies numbered and signed by the author. No. 169.”
 © Jan. 21, 1932; 2c. and aff. Jan. 25; A 47485; Doubleday, Doran & co., inc. (32-3525) 759

Source: *Catalogue of Copyright Entries, Part 1, Group 1: Books* for 1932.

TABLE 3
Example of a Transcription Entry

Type	Author	AuthDates	Edition	Title	Pubplace	Publisher	Pubdate	Copyright-date	Receipt-date	Reg-ID	Claimant	LCNN	Part from Vendor	Fascicle Number	Spanning Year	Order Number
book	Huxley, Aldous Leonard	1894	First edition	Brave new world	Garden City, N.Y.	Doubleday, Doran & company, inc.	1932	Jan. 21, 1932	2c. and aff. Jan. 25	A 47485	Doubleday Doran & co., inc.	32-3525	1932_Par 11	no. 8, March 26, 1932	1932 to March 29, 1933)	15205

Source: Transcription of the *Catalogue of Copyright Entries, Part 1, Group 1: Books* for 1932.

A circular released by the Copyright Office in 1927 describes the contents of *PIG1: Books* as “all important books, both domestic and foreign, deposited and registered for American copyright.”²⁸ The term “important” here is vague, and unfortunately the Copyright Office fails to clarify how it decided which books were “important.”

From these observations about *PIG1: Books*, it is reasonable to conclude that there are more books than those listed in *Part 1, Group 1: Books*. Furthermore, the Copyright Office’s annual reports indicate that the number of “books proper” registered is significantly higher than the number of entries in the “important book” volume: 13,460 for FY1932 (July 1931 through June 1932), and 10,820 for FY1933 (July 1932 through June 1933).²⁹ Therefore, we extended our analysis of books published in 1932 to other parts of the *CCE*, recognizing that there may be more books than the 9,004 “important” ones in *PIG1: Books*.

Part 1, Group 2: Pamphlets, Contributions to Newspapers or Periodicals, Lectures, and Maps

Part 1, Group 2 (PIG2: Pamphlets) contains most of the Class A copyrights.³⁰ The Copyright Office’s *Circular no. 36* from 1927 described this part as “Pamphlets, dramatic compositions, maps and motion pictures, contains titles of books and pamphlets not included in Group 1.”³¹ *PIG2: Pamphlets* contains registrations for Class A and a subtype called Class AA, as well as some foreign and ad interim copyrights. Similar to the definition of “important,” a definition of the AA subtype does not appear in the *CCE* until a later volume, although the class itself first appeared in 1929.³² The 1947 *CCE* describes Class AA as “pamphlets, leaflets, etc.” and “serials registered in the Copyright Office.”³³ The choice to use A rather than AA is not explained, nor is it evident from examining the entries.

Fascicles of *PIG2: Pamphlets* were published monthly, which was less frequent than *PIG1: Books*. Each monthly fascicle contained a summary table at the end. The table

TABLE 4 1932 Number of Entries in Part 1 Group 2: Pamphlets of the Catalogue of Copyright Entries	
Type	Entries
(a) United States publications including 10,160 contributions to periodicals and other 1-copy entries	38,398
(b) Foreign books in foreign language	1,594
Total	39,992
<i>Source: Catalogue of Copyright Entries, Part 1, Group 2: Pamphlets for 1932.</i>	

rolled over totals from previous months, so that the final summary table reflects the whole volume (see table 4).³⁴ As can be seen in the example in figure 3, entries for the *PIG2: Pamphlets* are briefer than *PIG1: Books*, are printed in two columns, and contain minimal bibliographic information.³⁵ Entries exclude the size of the materials and rarely include publisher information. Therefore, the nature of the works in the *PIG2: Pamphlets* volume is broad and ambiguous.

FIGURE 3 Example of 1932 PIG2: Pamphlets Entries	
Ponca City, Okl. Polk's Ponca City, Okl., directory, 1932. © Mar. 1; A 49141; R. L. Polk & co. 10370	
Ponton (W. S.) inc.* Ponton's 1932 mailing list catalog. © Mar. 15, 1932; AA 92757. 10371	
Poor's pub. co.* Poor's analytical service. © Mar. 24, 1932; A 49387. 10372	
Poppen (H. Y.)* What the contact exchange system, etc. © Feb. 8, 1932; AA 90941. 10373	
<i>Source: Catalogue of Copyright Entries, Part 1, Group 1, Pamphlets etc. for 1932</i>	

The words “books” and “pamphlets” are often used interchangeably in Copyright Office publications to describe the contents of this volume. For example, the title page of *PIG2: Pamphlets* refers to itself as *Part 1: Books, Group 2*, while a table on page ii of the same volume refers to it as *Part 1, Group 2: Pamphlets, Contributions to Newspapers or Periodicals, Lectures, and Maps*.³⁶ It is important to remember that, despite separate books and pamphlets volumes of the CCE publication, the actual entries were legally

classified the same (Class A), pursuant to the Copyright Act of 1909. The use of this class for both volumes likely accounts for why a broad swath of formats (that is, books, directories, leaflets) would all confusingly be called “books.” “Important books,” versus those that were not, appears to matter only for the CCE publication and has no bearing on the actual status of the copyright itself. A copyright in *P1G1: Books* volume has the same value as one in *P1G2: Pamphlets*. Splitting Class A into different volumes is a way to organize the entries so books more likely to need copyright protection (those that came from the publishing trade) are published in the CCE within a few days of arriving at the Copyright Office.

As previously noted, *P1G2: Pamphlets* is described as including books. There are 13,460 “books proper” registered according to the FY1932 *Annual Report*, 10,820 in the FY1933 *Annual Report*, but only 9,004 “important” books in the 1932 CCE. An explanation for the discrepancy is that there are “books proper” recorded in the *P1G2: Pamphlets* volume, too. This is still somewhat speculative, since the Copyright Office fails to define or classify “books proper” anywhere. To support this hypothesis further, a sampling of the 39,992 *P1G2: Pamphlets* entries was examined to better understand

TABLE 5
Sampling of Formats from the Pamphlet Volume of the Catalogue of Copyright Entries

Format	Total Titles
Advertisement	252
Product Manual	216
Article in Journal/Newspaper	120
Product Catalog	86
Book	75
City Directory	58
Telephone Directory	42
Information Sheet	36
Report	15
Workbook	11
Yearbook	11
Juvenile book	10
Guidebook	9
Poetry	8
Handbook	7
Lecture	7
Textbook	7
Game Instructions	6
Teacher’s guide	6

TABLE 5
Sampling of Formats from the Pamphlet Volume of the Catalogue of Copyright Entries

Format	Total Titles
Bulletin	5
Lab manual	5
Sunday cartoon sheets	5
Advertisement Card	4
Circular	4
Cook book	4
Exhibition Catalog	4
Answer book	3
Chapter in Book	3
Chart	3
Coloring Book	3
Citations	2
Field Guide	2
Instructor’s guide	2
Inventory	2
Outline	2
Pamphlet	2
Calendar	1
Grand Total	1,038

whether books could also be included in this volume. The sample size was 1,038 entries from the *P1G2: Pamphlets* volume, using a 95 percent confidence level and a confidence interval of 3.³⁷ Based on the limited data in the entry, each was ascribed a bibliographic form, such as advertisement or product manual. Rather than begin with a master list of forms (such as Library of Congress Genre Terms), they were determined based on the given information in the entries (see table 5).³⁸

When sampling the *P1G2: Pamphlets* volume, we found that there were many materials that obviously could not be books, such as advertisements and journal or newspaper articles. Even so, entries were also found that could be books, such as juvenile books, self-published poetry, handbooks, textbooks, and cookbooks. Without further data, there is still a lot of uncertainty regarding what most of these materials actually were. In our sample, we found 111 entries in just these book-like categories (juvenile books, self-published poetry, handbooks, textbooks, and cookbooks); by extrapolating that number against the total number of entries, we get an estimate of 4,106 books in the *P1G2: Pamphlets* volume. The gap between “important” books and “books proper” was 4,456, so our sample estimate of possible books accounts for 92 percent of the gap for the 13,460 books noted in the *Annual Report* for FY1932. This supports our hypothesis that the Copyright Office put some books in *P1G2: Pamphlets*.

To be clear, the Copyright Office does not explicitly note how many books were entered into the *P1G2: Pamphlets* volume; however, a few thousand books in this volume would account for the discrepancy between the number of “books proper” tallied in the *Annual Report* and the number of “important” book entries in *P1G1: Books*. From this, we can infer that the Copyright Office received material that they would describe as a “book” but decided it was not important enough for *P1G1: Books* so listed the entries in *P1G2: Pamphlets*. The *P1G1: Books* volume was published on a rigorous schedule, so it is not difficult to imagine certain works getting prioritized over others due to their publisher or author. Therefore, this reinforces a suspicion that the first volume of the *CCE: Part 1, Group 1: Books* (important books) is not a fully comprehensive list of all the books copyrighted in 1932.

Evaluating Ringer’s “Study 31: Renewal of Copyright”

To assess the number of renewals for books only, Ringer’s study must be recreated with more precise data on how many books were included in the copyright registrations and renewals. Ringer used the total number of book registrations listed in the FY1932 *Annual Report* for Class A (57,065) (see table 6) as her number of registrations, and the total number of renewals from the July 1958 through June 1959 *CCEs* (3,942).³⁹ From these sources, Ringer determined the renewal rate for Class A materials was 7 percent.⁴⁰ As Wilkin noted, those copyrights include thousands of nonbook materials.⁴¹ She did note in her report that Class A “Includes contributions to periodicals,” but this insufficiently describes just how many of Class A materials were not actually books. There were 26,995 pamphlets and 10,489 periodical contributions, along with an assortment of ad interim copyrights and foreign language publications included, which inflates the total number of Class A registrations.

Fortunately, there are a few ways of narrowing down Class A materials to only books. Our transcription of the *CCE* volume, *P1G1: Books* lists 9,004 “important” books for calendar year 1932. The *Annual Report* also provides the number of “books proper” for FY1932–FY1933: 13,480 and 10,820.⁴² Even though there are multiple possible totals for books (9,004, 10,820, and 13,460) they are all significantly lower than Ringer’s 57,065 and provide more accurate options for book renewal rate calculation.

During the analysis, duplicated registration IDs were observed in the data. The following reasons explain why duplicated registration IDs appear in the *CCE* transcription:

TABLE 6
FY 1932-1933 Registrations for Class A from the Copyright Office *Annual Report*

Books:		FY1932	FY1933
(a) Printed in the United States—			
	Books proper	13,460	10,820
	Pamphlets, leaflets, etc.	26,995	24,290
	Contributions to newspapers and periodicals	10,489	9,290
	Total	50,944	44,400
(b) Printed abroad in a foreign language		4,784	4,232
(c) English books registered for ad interim copyright		1,337	1,352
Total		57,065	49,984
Source: Copyright Office’s <i>Annual Report</i> for 1932 and 1933.			

- Some titles appear multiple times in the CCE with identical entries, such as *Wanda Gág’s Story Book*. This is likely a clerical error, so the duplicate registration was removed.
- A registration ID is used twice (arbitrarily) for two unrelated works. An example of this is *Sheriff Killer* by Dane Coolidge, which has the same registration ID as *Your Money* by Samuel Crowther. The books have no metadata in common to suggest a link between the two; the Copyright Office just reused the registration ID. In these instances, we created a dummy number so that each registration had a unique registration ID.
- There are also instances where multiple works by the same author or publisher shared a registration ID. For example, a series of histories by P.F. Collier & Son Company (*England; Ireland; The French Revolution from 1789 to 1815*) were all registered as A-49817. This was the most common reason for the duplication. Duplicated records for these titles were deleted so there is only one renewal per registration.

After the data of registrations were cleaned up, there are 8,967 unique registration IDs for 9,004 entries.

Instead of using the *Annual Report’s* number for total renewals, we use renewal data from the Copyright Renewal Database, created by Stanford University Libraries. The Copyright Renewal Database is a transcription of the *Catalogue of Copyright Entries* renewals for Class A works. A list of renewals for materials registered was produced by searching the Copyright Renewal Database for all renewals for 1931, 1932, and 1933 (typically made between 1958 and 1960) and exporting the search results as a table. Rather than using the raw number from the CCE, this method allows us to have individual records of each renewal.

Unfortunately, Stanford’s database is missing some data. The database has each renewal but does not include every registration ID or registration date listed in the original entries. For example, the *Encyclopedia of Automobile Law, Vol. 3–4, 5–6. 9th ed.* by Xenophon Huddy has two registration IDs, A-32663 (made on January 7, 1931) and A-33548 (made on February 2, 1931) and two renewal numbers, R223321 and R223322.⁴³ Stanford’s database lists both renewal IDs, but only the first date and registration ID: A-32663. This makes it appear as though there are two renewals (R223321 and R223322) for one registration ID (A-32663). Therefore, the Stanford database does not list every

registration ID, even when it lists every renewal. Not all entries fail to list every registration ID and date for multipart works, but many do. This will likely underrepresent the number of renewals there are when we link registration IDs and renewals.

The data downloaded from Stanford's Copyright Renewal Database also requires a lot of cleanup before it can be used to analyze renewal rates. The following are the issues with the copyright renewal data and how it was cleaned up for this project:

- Many Class B renewals were included in the Renewal Database, which were removed from our list along with foreign and ad interim copyrights.
- Obvious articles from serial publications were removed from the list, such as registration IDs that begin with A5- or works described as "in" other publications (for instance, "C. I. D." [In *Adventure magazine*]).
- Multiple registration IDs were listed in a single cell. These were cleaned up so each row of the spreadsheet has a single registration ID.
- A single title with one registration was renewed more than once for the same 28-year period. This happened with *Little House in the Woods*, which was renewed three times for a single registration: twice by Rose Wilder Lane (for illustrations and text) and once by the publisher, Harper & Bros. This was deduplicated to retain a one-to-one relationship between registrations and renewals.

If uncorrected, cases of duplicated registration IDs inflate the number of renewals. For example, for 1932 after the duplicated registration IDs were removed, the number of renewals on the list for that year declined from 3,791 to 3,247.

Renewal Rate Calculations

After analyzing the different sources of copyright registration and renewal data, we decided there should be two approaches to calculating a renewal rate for books. They will be referred to as the Important Books calculation and the Books Proper calculation, borrowing the Copyright Office's terminology.

Important Books Calculation

The first option is to use the Copyright Office's decision about which books are "important": those listed in our transcription of the *Part 1, Group 1: Books*. This assumes, based on the Copyright Office's CCE publication, that the number of books registered for copyright for 1932 is 9,004. After the data are cleaned up and a few registration IDs were deduplicated, the total number of unique registration IDs is 8,967.

Using the Copyright Renewal Database renewals for 1932, the data were matched against the *PIG1: Books* transcription to see how many "important" books were renewed. The data sets were matched by copyright registration ID.⁴⁴ The report determined that 2,489 registration IDs (from the renewal list) match 8,967 "important" book registration IDs. This indicates that 2,489 "important" books registered in 1932 were renewed out of the 8,967 total. Under this calculation, the renewal rate for "important" books is 28 percent.

The first issue with this calculation is that neither the Stanford database nor our transcription of 1932 CCE entries lists every registration ID, so some matches may have been missed; this would underreport the number of renewals. There is also a flaw in this method: it relies on the Copyright Office's designation of "important book." There were renewals that did not match any registration IDs in this dataset, because the renewal came from a source other than the "important book" (*PIG1: Books*) volume. This means that, while the Copyright Office decided materials did not qualify as "important" books, they were important enough to the copyright holder to renew 28 years after publication (when the copyright was eligible for renewal). There were 758 registrations renewed for copyright, presumably from the pamphlet volume. This is

758 renewals for 38,398 original registrations.⁴⁵ Therefore, a small number of Class A material from the *P1G1: Pamphlets* volume were renewed.

Books Proper

According to their *Annual Report*, the Copyright Office registered 13,460 “books proper” in FY1932. Since the annual reports use the fiscal year from July to June, the number of “books proper” must be taken from fiscal years 1932 and 1933 to overlap with calendar year 1932. These numbers are 13,460 (FY1932) and 10,820 (FY1933).⁴⁶

Using the annual reports’ “books proper” totals (10,820 and 13,460) is appealing because the numbers are higher than the “important books” amount (9,004). Since the number of books is higher, this may include all books, even those excluded from the “important book” volume of the *CCE: Part 1, Group 1: Books*. These were the books that were instead located in *Part 1, Group 2: Pamphlets*. This amount can be compared to the total number of renewals from FY1932 (3,533) and FY1933 (3,612) from the Copyright Renewal Database. The renewal rate calculation for “books proper” is between 26 (3,533/13,460) and 33 percent (3,612/10,820) (see table 7).

Form	Year	Registrations	Renewals	Percent Renewed
Books Proper	FY 1932	13,460	3,533	26%
Books Proper	FY 1933	10,820	3,612	33%
Important Books	1932	9,004	2,489	28%

Fortunately, both renewal rate calculations, 28 percent from the *CCE* and 26 to 33 percent from the annual reports, are similar and reflect a substantial increase from Ringer’s 7 percent.

Conclusion

Overall, this project underlines the need for all records of United States copyright data to be digitized and made available to the public, especially the Copyright Card Catalog. While this study determined that the *CCE* can be reasonably trusted as a substitute for the Card Catalog, it also relies on sampling transcribed data. Due to the serial nature of the *CCE*, and the wide variation in binding practices and digitization procedures across libraries, the information in one copy of the *CCE* may differ from another. There are copies online of the *CCE* in HathiTrust and Google Books, and, because they are drawn from different print collections, they often differ in their online manifestation. For example, there is a digitized version of *Part 1, Group 1: Books* (new series, volume 2, 1907), from Cornell’s⁴⁷ copy attached to one HathiTrust record, and a different digitized version of the same volume from the University of Michigan⁴⁸ attached to a different record. Transcriptions would almost certainly contain deviations between copies in both physical and digital libraries due to the Copyright Office releasing them as weekly fascicles (in other words, Michigan’s copy is 1,343 pages while Cornell’s is 1,139 pages, indicating a difference in content between the two versions). Considering these issues, the best source of data for analyzing copyright registrations and renewals is the original data held in the Copyright Card Catalog, which contains information on copyright registrations and renewals from 1870 to 1977.⁴⁹ At this time, the only way to access pre-1978 data is to either use the *CCEs* or visit the Copyright Office Reading Room in person. There was a digitization initiative to scan the cards for preservation

that was completed in September 2014. These data will be added to the online catalog after the metadata from each card is “reviewed and data perfected.”⁵⁰ However, no indication of the expected timeline for this is indicated. Currently, the online catalog only contains registrations for 1978 up to the present day.

While the online catalog is a step in the right direction, several issues need to be addressed. At present, the 1978 catalog does not support batch searching (or searching for more than one registration at a time). Additionally, the data it holds are not exportable for use outside of the Copyright Office’s system, severely restricting opportunities to mine or enhance the data. In this data-driven age, it is imperative that not only should the digitized pre-1978 data be made available online as soon as possible, all of the data from 1870 to present day should be published in a machine-readable format that enables use by researchers outside of the Copyright Office’s environment.

Wilkin’s article was correct to critique the assumption that 7 percent of Class A registration renewals reflected the number of books renewed for 1932. He was also correct in hypothesizing that the renewal rate would be much higher than 7 percent. In his paper, Wilkin also questioned the premise that there were 13,460 books published in 1932, because the *CCE* only noted 8,994 entries for the same year, and roughly the same amount for the years before and after 1932. Wilkin was incorrect in his assumption that we should necessarily determine the copyright renewal rate from the *PIG1: Books* volume (9,004 registrations) rather than the *Annual Report* (13,460 registrations). This study determined that it is likely that the Copyright Office deposited book registrations in the pamphlet volume, making the range of 10,820–13,460 a more accurate reflection of all books published for 1932, regardless of their importance. Consequently, we believe the range of 26 to 33 percent for a renewal rate based on “books proper” registrations is more accurate. The “important book” calculation falls within this range supporting this conclusion.

Given the inaccuracies and inconsistencies in the copyright registration and renewal data, this study calls for a reexamination of the conclusion made by Demas and Brogdon in “Determining Copyright Status for Preservation and Access: Defining Reasonable Effort.”⁵¹ The study concluded that the *CCE* is 97 percent accurate on registrations and 100 percent accurate on renewals, and therefore authoritative on copyright status. We confirm that the *CCE* is a reasonable substitute for the Card Catalog (the original registrations held at the Library of Congress). Demas and Brogdon found that only 3 out of 370 titles in their study were in the *PIG2: Pamphlets* volume. They appeared in the renewals, so this was not a hindrance to discovery of copyright status. The extensive analysis done in this study of the *CCE* adds a caveat to that conclusion. Since this study determined that thousands of books are likely in the *PIG2: Pamphlets* volume, it is necessary to search this volume as well when making copyright status determinations, even if the item is a book.

Limiting the renewal rate calculation to just books is important because of the work being done to digitize library collections. Out of the 9,004 registrations in the *PIG1: Books* volume, 8,990 (99.8%) could be linked to titles in OCLC WorldCat or the Library of Congress catalog, meaning that these works should be present in at least one OCLC WorldCat member library and/or the Library of Congress. Therefore, “important books” often correlates with library holdings, which are also the books being selected for digitization by HathiTrust and Google Books. It is crucially important to have an accurate renewal rate for these materials. A renewal rate of 26 to 33 percent for FY1932 is significantly higher than 7 percent, and this rate would increase over time; that conclusion necessitates a shift in how we understand copyright renewal for books of the mid-twentieth century. For libraries and our stakeholders in particular, this means that, as we digitize our collections, we must understand how much of our collections still remain protected by copyright.

Notes

1. Barbara Ringer, "Study No. 31: Renewal of Copyright," reprinted in *Library of Congress Copyright Office, Copyright Law Revision: Studies Prepared for the Subcommittee on Patents, Trademarks, and Copyrights of the Committee on the Judiciary, United States Senate, Eighty-sixth Congress, First [Second] Session* (Washington, D.C.: Government Printing Office, 1961), 109.
2. Senate and House of Representatives of the United States of America in Congress assembled, *Copyright Law of the United States, Act of March 4, 1909 (in effect July 1, 1909)*, 60th Cong., 2d sess., 1909, 3, available online at <https://www.copyright.gov/history/1909act.pdf> [accessed 19 April 2017].
3. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1 Group 1 Books Including List of Renewals: New Series Volume 29, Group 1, Volume Index for the Year 1932* (Washington, D.C., 1933), ii, available online at <http://hdl.handle.net/2027/mdp.39015076106494> (Source of total book registrations) [accessed 19 April 2017]; Stanford University Libraries, *Copyright Renewal Database* (2004), distributed by Stanford University Libraries, available online at <https://collections.stanford.edu/copyrightrenewals/bin/page?forward=home> (Source of total book renewals) [accessed 5 January 2017]; Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1 Books Group 2: New Series Volume 29, Part 2, Last Half of 1932, Nos. 9–12 with Annual Index* (Washington, D.C., 1933), 1121, available online at <http://hdl.handle.net/2027/mdp.39015076106874> (Source of total pamphlet registrations) [accessed 17 April 2017].
4. John P. Wilkin, "How Large Is the 'Public Domain'? A Comparative Analysis of Ringer's 1961 Copyright Renewal Study and HathiTrust CRMS Data," *College & Research Libraries* 78, no. 2 (2017): 207, available online at <https://doi.org/10.5860/crl.78.2.201> [accessed 17 March 2017].
5. The CRMS project is documented online at www.lib.umich.edu/copyright-review-management-system-imals-national-leadership-grant [accessed 12 December 2016]. A full description of the project is included there.
6. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1. Group 1 Books Including List of Renewals: New Series Volume 29, Group 1, Volume Index for the Year 1932*, 120.
7. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1 Books Group 2: New Series Volume 29, Part 2, Last Half of 1932, Nos. 9–12 with Annual Index*, 897.
8. Ringer, "Study No. 31: Renewal of Copyright," 221.
9. Peter B. Hirtle, "When Is 1923 Going to Arrive and Other Complications of the U.S. Public Domain," *Searcher* 20, no. 7 (2012), available online at <http://www.infotoday.com/searcher/sep12/Hirtle-When-Is-1923-Going-to-Arrive-and-Other-Complications-of-the-U.S.-Public-Domain.shtml> [accessed 23 May 2018].
10. Library of Congress, Copyright Office, *Thirty-fifth Annual Report of the Register of Copyrights for the Fiscal Year Ending June 30, 1932* (Washington, D.C., 1932), 17, available online at <https://www.copyright.gov/reports/annual/archive/ar-1932.pdf> [accessed 17 April 2017].
11. United States Copyright Office, "Overview of the Copyright Office," available online at www.copyright.gov/about/ [accessed 19 April 2017].
12. Thomas G. Tanselle, *Selected Studies in Bibliography* (Charlottesville: Published for the Bibliographical Society of the University of Virginia by the University Press of Virginia, 1979), 103.
13. Joseph William Rogers, *U.S. National Bibliography and the Copyright Law: An Historical Study* (New York: R.R. Bowker Company, 1960), 67–68, available online at <http://hdl.handle.net/2027/mdp.39015030337730> [accessed 19 April 2017].
14. Copyright Office, Library of Congress, *Sixty-second Annual Report of the Register of Copyrights for the Fiscal Year Ending June 30, 1959* (Washington, D.C., 1960), 3, available online at www.copyright.gov/reports/annual/archive/ar-1959.pdf [accessed 19 April 2017].
15. Library of Congress, Copyright Office, *Thirty-fifth Annual Report of the Register of Copyrights for the Fiscal Year Ending June 30, 1932*, 17.
16. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1 Group 1 Books Including List of Renewals: New Series Volume 29, Group 1, Volume Index for the Year 1932*, iv.
17. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1 Group 3 Dramatic Compositions and Motion Pictures: Volume 5 for the Year 1932, Nos. 1–12* (Washington, D.C., 1933), 1, available online at <http://hdl.handle.net/2027/mdp.39015076106742> [accessed 19 April 2017].
18. Library of Congress Copyright Office, *Catalogue of Copyright Entries Part 3 Musical Compositions: New Series Volume 27, No. 1, First Half of 1932 Nos. 1–6* (Washington, D.C., 1932), 10, available online at <http://hdl.handle.net/2027/mdp.39015077982208> [accessed 19 April 2017]; Copyright Office, the Library of Congress, *Catalog of Copyright Entries Third Series Volume 1, Parts 3–4, Number 1: Dramas and Works Prepared for Oral Delivery, January–June 1947* (Washington, D.C., 1947), available online at <http://hdl.handle.net/2027/mdp.39015085477217> [accessed 19 April 2017] (P1G3: *Dramatic Compositions* and P3: *Musical Compositions* were both published monthly. Most

dramatic compositions were given a prefix D and musical compositions were given prefix E, and thus will not be discussed here. However, there are occasional Class A and AA registrations in both volumes. Apparently, in 1932, the Copyright Office chose to place some Class A registrations according to their genre rather than their class. By 1947 there are no longer class A or AA entries with dramatic or musical compositions. The number of Class A entries from these volumes is too minimal to impact renewal rate calculations in a significant way. As was stated earlier, there are Class A entries in *P1G3: Dramatic Compositions* and *P3: Musical Compositions* that duplicate entries in *P1G1: Books* or *P1G2: Pamphlets*, while others are unique. Although the presence of Class A/AA materials in Drama and Music volumes is small, they are still represented in the renewals for material published in 1932, so it is worth noting.); A sampling was taken of the *P1G3: Dramatic Compositions* volume to determine how often Class A copyrights were entered here. The sample was 367 (95% confidence level, confidence interval 5) from a total 8,024 entries in the volume. There were 5 Class A entries and 1 Class A ad interim, which is 1.36 percent. All entries from this sample were duplicated in *P1G1: Books* in 1931 or 1932.

19. Library of Congress, Copyright Office, *Catalogue of Copyright Entries, Part 1, Group 1, Books Including List of Renewals: New Series Volume 29, Group 1, Volume Index for the Year 1932*, 1813.

20. Library of Congress, Copyright Office, *Catalogue of Copyright Entries, Part 1, Group 3, Dramatic Compositions and Motion Pictures: Volume 5 for the Year 1932 Nos. 1–12*, 362.

21. Some Class A works share a registration ID without sharing a title or author information that might imply a link between the two. For example, *The Life Story of Greta Garbo* by Mary Margaret McBride and *Catalogue of the Collection* by the Whitney Museum of American Art are both registered as A-49350.

22. Ad interim copyrights are works published in the English language abroad. It holds the copyright “ad interim” until the item is published in the United States. Therefore, there are usually two registrations for the work: the first interim registration for the foreign publication, followed by the standard copyright for when it is published in the United States.

23. The vendor used was Apex CoVantage, 198 Van Buren Street, 200 Presidents Plaza, Herndon, VA 20170-5338, <http://apexcovantage.com/>. They took approximately three weeks to transcribe three years of the CCE.

24. Library of Congress, Copyright Office, *Catalogue of Copyright Entries, Part 1, Group 1, Books Including List of Renewals: New Series Volume 29, Group 1, Volume Index for the Year 1932*, iii.

25. *Ibid.*, ii.

26. The reason for the discrepancy is unknown. It could likely be a clerical error that happened when the Copyright Office counted entries.

27. Library of Congress, Copyright Office, *Catalogue of Copyright Entries, Part 1, Group 1, Books Including List of Renewals: New Series Volume 29, Group 1, Volume Index for the Year 1932*, 120.

28. Library of Congress, Copyright Office, *Catalogue of Copyright Entries [on Content and Distribution]: Explanatory Circ. No. 36* (Washington, D.C., 1927), available online at <http://congressional.proquest.com:80/congressional/docview/t66.d71.lc3.4-129.3?accountid=14553> [accessed 19 April 2017].

29. Library of Congress, Copyright Office. *Thirty-sixth Annual Report of the Register of Copyrights for the Fiscal Year Ending June 30, 1933* (Washington, D.C., 1933), 18, available online at www.copyright.gov/reports/annual/archive/ar-1933.pdf [accessed 19 April 2017].

30. Registrations for maps and lectures were also bound into this volume, although they are in their own sections and have separate prefixes. Therefore, they are not discussed as part of the analysis.

31. Library of Congress, Copyright Office, *Catalogue of Copyright Entries [on Content and Distribution]: Explanatory Circ. No. 36* [emphasis ours].

32. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Published by Authority of the Acts of Congress of March 3, 1891 of June 30, 1906, and of March 4, 1909: Part 1 Books Group 2: Pamphlets, Leaflets, Contributions to Newspapers or Periodicals, etc.; Lectures, Sermons, Addresses for Oral Delivery; Maps, 1929, New Series, Volume 26, Part 1* (Washington, D.C., 1929), 1, available online at <http://hdl.handle.net/2027/mdp.39015076106932> [accessed 19 April 2017].

33. Library of Congress, Copyright Office, *Catalog of Copyright Entries, Third Series, Volume 1, Part 1B, Number 1: Pamphlets, Serials and Contributions to Periodicals January–June 1947* (Washington, D.C., 1947), iii, available online at <http://hdl.handle.net/2027/mdp.39015085477944> [accessed 19 April 2017].

34. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1 Books Group 2: New Series Volume 29, Part 2, Last Half of 1932, Nos. 9–12 with Annual Index*, 1121.

35. *Ibid.*, 342.

36. *Ibid.*, ii.

37. Creative Research Solutions, “Sample Size Calculator,” available online at www.survey-system.com/sscalc.htm [accessed 17 March 2017].

38. The Library of Congress Form/Genre terms do not include many terms that are necessary to categorize P1G2. For example, there is no genre/form term for “advertisements” or “articles.”

39. Library of Congress, Copyright Office, *Catalog of Copyright Entries Third Series Volume 12, Part 1, Number 2: Books and Pamphlets Including Serials and Contributions to Periodicals July–December 1958* (Washington, D.C., 1959), xi, available online at https://books.google.com/books?id=MR4hAQAAlAAJ&pg=PR3&source=gbs_selected_pages&cad=2#v=onepage&q&f=false [accessed 19 April 2017]; Library of Congress, Copyright Office, *Catalog of Copyright Entries Third Series Volume 13, Part 1, Number 1: Books and Pamphlets Including Serials and Contributions to Periodicals January–June 1959* (Washington, D.C., 1960), xi, available online at https://books.google.com/books?id=TSEhAQAAlAAJ&pg=PR3&source=gbs_selected_pages&cad=2#v=onepage&q&f=false [accessed 19 April 2017].

40. Ringer, “Study No. 31: Renewal of Copyright,” 221.

41. Wilkin, “How Large is the ‘Public Domain?’” 204.

42. Ringer could not use the “books proper” numbers because she had no way of knowing how many renewals made 28 years later came from total number of “books proper” registrations, since she only had summary totals for all of Class A renewals (not a more granular number of renewals for just books proper).

43. Library of Congress, Copyright Office, *Catalog of Copyright Entries Third Series Volume 12, Part 1, Number 2: Books and Pamphlets Including Serials and Contributions to Periodicals July–December 1958*, 1623.

44. After the transcription and renewal lists were compared using the registration ID as the link, the results were checked for any possible mismatches. It was determined that 7 copyright renewals incorrectly matched with registrations. The cause was the reuse of those registration IDs in the original 1932 P1G2: *Pamphlets* volume. Those false matches were deleted from the final result.

45. Since there is no transcription of the P1G2: *Pamphlets* volume, an assumption is being made here that if the renewal did not match a registration, it belonged in another volume: pamphlet, drama, or music. However, a failure to match just means the renewal was not registered initially in P1G1: *Books*. It is possible the match failed because the transcription did not correctly input the registration ID, or it was registered another year.

46. Library of Congress, Copyright Office. *Thirty-sixth Annual Report of the Register of Copyrights for the Fiscal Year Ending June 30, 1933*, 18.

47. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1: Books; Dramatic Compositions; Maps and Charts; Index New Series, Volume 2, January–June 1907, Nos. 1–26*, (Washington, D.C., 1907), available online at <https://hdl.handle.net/2027/coo.31924112597350> [accessed 19 September 2017].

48. Library of Congress, Copyright Office, *Catalogue of Copyright Entries Part 1: Books; Dramatic Compositions; Maps and Charts; Index New Series, Volume 2, January–June 1907, Nos. 1–26*, (Washington, D.C., 1907), available online at <https://hdl.handle.net/2027/mdp.39015076105991> [accessed 19 September 2017].

49. Library of Congress, Copyright Office, *The Copyright Card Catalog and the Online Files of the Copyright Office*, available online at <https://www.copyright.gov/circs/circ23.pdf> [accessed 19 September 2017].

50. “Project Status,” available online at <https://www.copyright.gov/digitization/status.html> [accessed 19 September 2017].

51. Samuel Demas and Jennie L. Brogdon, “Determining Copyright Status for Preservation and Access: Defining Reasonable Effort,” *Library Resources & Technical Services* 41, no. 4 (Oct. 1997): 329, doi:10.5860/lrts.41n4.323.