TITLE: Case Study on the Book Republication Program in 1942: Examine the Effects of Copyrights on Science

AUTHOR: Yuan Huang (yh2741@nyu.edu)

DATE SUBMITTED: December 2, 2020

INTRODUCTION

Copyright establishes intellectual property in science, and it captures the economic value of the science work, which can increase the incentive of scientists to develop new knowledge. However, on the other hand, the copyright laws create high access costs for researchers which impose a deadweight loss for future innovations, especially for science innovation which develop critically on access to existing knowledge.

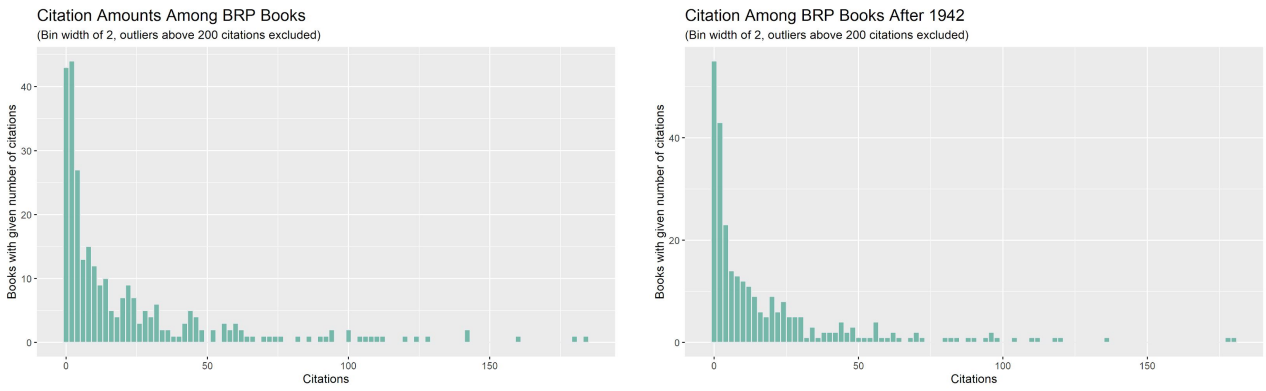
This report exploits an exogenous change in copyrights to examine how the Book Republication Program influenced the creation of new knowledge of science by reducing the price of German-owned science books. In 1942, the U.S. government established the Book Republication Program, and it allowed the U.S. publishers to reprint the exact content of German-owned science books. Before the implementation of the Book Republication Program, German-owned copyrights were protected as the same as American books by the 1909 Copyright Act (Moser & Biasi, 2018. p. 5). By increasing the supply of reprinted German-owned science books, the price of books decreased. The change of the prices allows us to figure out the effects of lower access costs for science books in two different copyright regimes.

The empirical analysis examines the data for the total number of 291 books subjected to BRP(Book Republication Program) and 486 books not subjected to BRP between 1920 to 1970. Among all the fields in BRP books, the field of compounds was most frequently licensed. The dataset is collected by Professor Petra Moser and Professor Barbara Biasi. In the datasets, it includes the book ID, year of citation of the book, language of the books, research fields of books, citation languages *etc*. In this report, we use future articles and books that cite BRP and non-BRP books as a measure for new knowledge that builds on those books.

EMPIRICAL EVIDENCE AND DISCUSSION

For the empirical analysis, we use two identification strategies to examine the effects of the BRP to the invention of new scientific knowledge. The two identification strategies aim to address two main challenges for identifying the causal effect of the BRP to the copyright on science. The first challenge is that the selection of BRP books was not random. To address this issue, we need to compare the change in citations from English-language authors who benefit more directly from this program with other-language authors who benefit less from the same BRP book. The second challenge is that authors publishing in other languages were hit harder by WWII than English-language authors. As we all know, Europe was destroyed by WWII so that English language citation may increase mechanically due to the increased research output compared with Europe. To address this problem, we need to compare the changes in citations by English-language authors to BRP books and Swiss books before and after 1942. Swiss is chosen as the control group due to its neutrality during WWII, and unlike the German-owned books, Swiss books were not available to apply to the BRP.

FIGURE 1. TOTAL NUMBER OF CITATIONS FOR BRP BOOKS (ALL YEAR AND AFTER 1942)



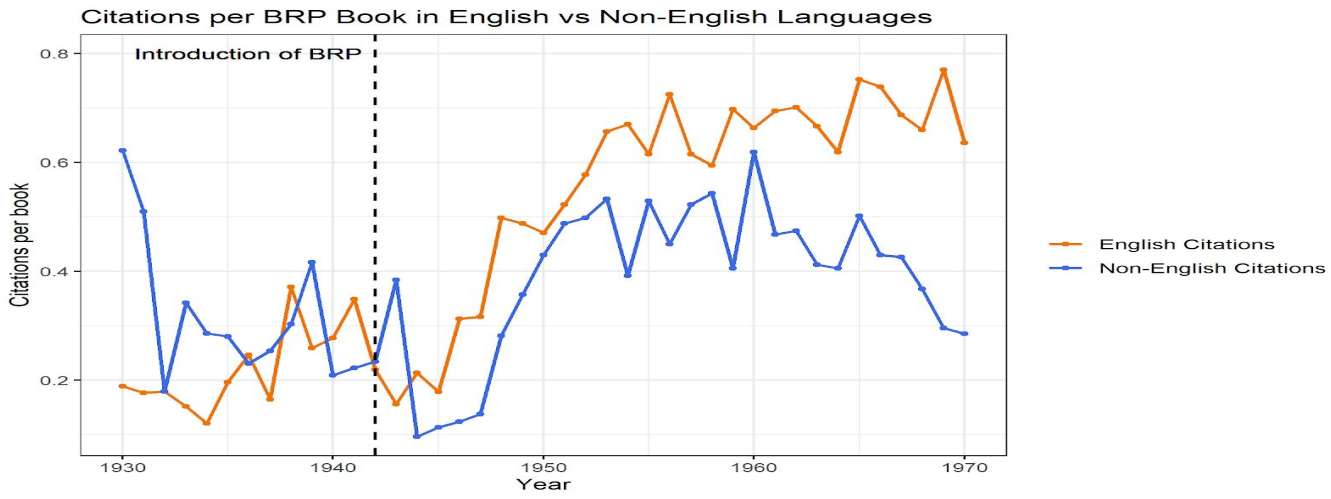
Notes: In these two histograms, we cut off the citations which are greater than 200 with the bin width equal to 5. On the left hand side, it is the citation among BRP books from 1920 to 1970. On the right hand side, it is the citation among BRP books from 1942 to 1970.

From Figure 1, we could see that both of the histograms exclude the outlier for citations larger than 200 and on the left-hand side is the total amount of citations for all years as for the right-hand side is the total amount of citations after 1942. By comparing the two graphs, we could find that the citations are similar between the two graphs. It may imply that most of the citations occurred after 1942. However, we need further empirical analysis to estimate the causal effect of the BRP.

Part I: The first Identification strategy

To avoid the selection bias for the BRP books, we need to compare the citations to the same BRP books by English-language authors and other-language authors. From Figure 2, we could find that before 1942, English citations and other language citations had similar trends which indicate that the parallel assumptions hold. The average of English citations per book is around 0.26, and the average of other language citations per book is almost 0.3. After 1942, we could observe that the English citations per book are above the other language citations per book with an average of 0.556 citations per book compared with 0.391 citations per book. It may imply that the introduction of BRP encourages the innovation of science.

FIGURE 2. CITATIONS TO BRP BOOKS IN ENGLISH CITATIONS AND OTHER LANGUAGE CITATIONS



Notes: English are citations to BRP books by English-language authors by the publication year of the citing publication. Other are citations from authors publishing in other languages). Citations are collected from Google Scholar (http://scholar.google.com, accessed July 1st-September 25th, 2014), and manually assigned to a publication language.

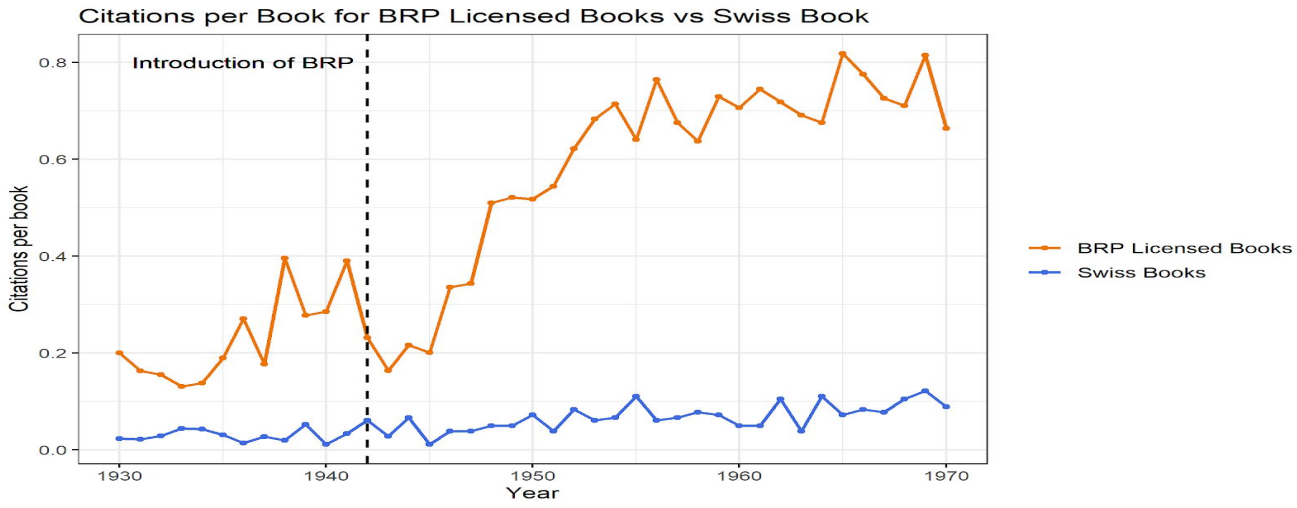
However, there is a potential issue for the identification strategy I -- the WWII because WWII led to a change among the control group and treated group at the same time. To fix this problem, we apply the second identification strategy and try to combine the first and the second identification to generate an appropriate difference-in-difference regression analysis.

Part II: The second Identification strategy

To avoid the influence of WWII that much of Europe was destroyed, we picked Swiss books as our control group, and we need to compare the changes in English-language citations for BRP books and Swiss books respectively before and after 1942. Furthermore, taking Identification Strategy I into consideration, we also need to pick a comparison group from all the books to

avoid the selection bias. By picking the proper comparison group, we first selected the field with the number of Swiss books larger or equal than 2 and then we selected the field in chemistry or mathematics. The reason is that chemistry and math were subjects that were licensed heavily since Germany led those fields, and many German authors published in the Swiss copyright system which would have exempted some of their work of comparably high quality from the BRP. The fields of chemistry and mathematics are also well represented in both the BRP group and the Swiss group. In this way, we focus on those two fields. We also dropped the dissertations since their behavior was quite different from an average science book, for instance, the open sources.

FIGURE 3. CITATIONS TO A SELECTED SAMPLE OF BRP AND SWISS BOOKS



Notes:Citations for a matched sample of 259 BRP books and 181 Swiss books.

In Figure 3, we could see that before 1942, the trend of new English citations to BRP books and Swiss books are similar, so that the parallel assumption holds under this situation. After 1942, it is evident that the new English citations for BRP books were much higher, which indicates that without copyright and lower the book price will encourage science. To systematically compare the changes in the English-language citation per book and year with and without copyright, we use the difference-in-differences regression in the following form:

$$cite_{it} = \alpha_t + \gamma_i + \beta_0 Post_t + \beta_1 BRP_i + \beta_2 Post_t \times BRP_i + \varepsilon_{it}$$

In the equation, the dependent variable $cite_{it}$ measure the new English citations to BRP and Swiss books under the book i and year t. $BRPi$ is 1 if the book is in the BRP and 0 otherwise. Also, α_t denotes as the time fixed effect, and γ_i denotes the book fixed effect.

TABLE 1. DIFFERENCE-IN-DIFFERENCES REGRESSION,
DEPENDENT VARIABLE IS ENGLISH CITATION PER BOOK AND YEAR, 1930--1970

	(1)	(2)	(3)	(4)
Treated X post-1942	0.277 ***	0.468 ***	0.523 ***	0.435 ***
	(0.047)	(0.025)	(0.022)	(0.034)
Post-1942	0.041			
	(0.032)			
Treated	0.247 ***			
	(0.042)			
Year FE	no	no	yes	yes
Book FE	no	yes	no	yes
N	16160	16160	16160	16160
R^2	0.043	0.552	0.049	0.560

*** p < 0.01; ** p < 0.05; * p < 0.1.

Notes: The dependent variable measures English-language citations to book i per year t between 1930 and 1970. The indicator BRP equals 1 for 259 books that were licensed to US publishers under the 1942 Book Republication Program (BRP). The control group includes 181 Swiss books that were not available for licensing due to Switzerland's neutrality during the war. The variable post equals to 1 for years after 1942.

From Table 1, difference-in-differences regression of equation (4) indicates that the BRP books had 0.435 additional new English citations per year and per book after 1942, and that is the treatment effect. The standard error is 0.034, and the p-value is smaller than 0.01 so that the estimate is statistically significant. By applying the year fixed effect and book fixed effect, we could make the estimate more accurate by treating different books and different years differently. With book fixed effects, we could control the influence caused by different types of knowledge.

CONCLUSION

This report uses an exogenous change in copyrights to examine how the Book Republication Program influenced the creation of new scientific knowledge. By using two identification strategies, we addressed the issues for selection bias and the influence of WWII, and we came up with a suitable comparison group to examine the effect of copyright on science innovation. Weaker copyright law will reduce the price of books and lower access costs, and it can encourage the creation of science. It is helpful for future policymaking that policymakers could encourage science innovation by lowering the access costs in the way of reducing the prices or applying subsidies, especially for the expensive textbooks.

Work Cited

Moser, Petra., & Biasi, Barbara. (2018, June 10). Effects of Copyrights on Science. *American Economic Journal.*