



Nakov, Novica (2016) *Human rights and pirate libraries: Can Sci-Hub.Org-enabled improvements in access to scientific literature help fulfil human rights?* [MRes.]

Copyright © 2016 The Author

Copyright and moral rights for this work are retained by the author(s)

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author(s)

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, institution and date must be given.

<http://endeavour.gla.ac.uk/135/>

Deposited: 9 December 2016

Enlighten Dissertations
<http://endeavour.gla.ac.uk/>
deposit@lib.gla.ac.uk

2201816

HUMAN RIGHTS AND PIRATE LIBRARIES:
CAN SCI-HUB.ORG-ENABLED IMPROVEMENTS IN
ACCESS TO SCIENTIFIC LITERATURE HELP FULFIL
HUMAN RIGHTS?

PRESENTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF

MRes in Human Rights & International Politics

UNIVERSITY OF GLASGOW

SEPTEMBER 2016

WORD COUNT: 13,440

Abstract

Pirate libraries offer universal and free access to scientific literature. They probably have every scientific article ever published in their repositories and have used human rights to justify their work. The extent to which they are used by researchers is largely unexplored. This dissertation evaluated the human rights claim of Sci-Hub.org, one of the leading pirate libraries, and investigated its usage by researchers in the European Union. Regressions were used to explore the relationships between downloads from Sci-Hub.org, and spending on academic libraries or published scientific articles, while taking into account internet access and prevalence of software piracy. Downloads from Sci-Hub.org were independent of spending on libraries but were influenced by both the number of published articles and internet access, suggesting that research needs influence usage of pirate libraries more than access needs. Widespread and continuously-increasing usage of Sci-Hub.org along with the independence from library spending supported pirate libraries as services that can help bridge the access gap for literature and more generally aid the advancement of human rights. However, universal availability of scientific literature, without research skills to complement it, could have limited effect on advancing human rights.

Table of Contents

1	Introduction	1
1.1	Aims, Objectives and Research Question	2
1.2	Summary of the Contents and Main Arguments of the Dissertation . .	3
1.3	Pirate Libraries	5
2	Literature Review	7
2.1	Introduction	7
2.2	Human Rights That Can Be Advanced by Better Access to Scientific Literature	8
2.2.1	The Right to Enjoy the Benefits of Scientific Progress	9
2.2.2	The Right to Information	11
2.3	Article Sharing as Human Rights Activism	14
2.3.1	Pirate Libraries as a Human Rights Struggle	15
2.3.2	Motivations for Contribution to Pirate Libraries	16
2.3.3	Critiques Against Pirate Libraries	19
2.4	The Value and Benefits of Access	22
2.5	Conclusion	24

3	Analysis of SciHub Data	27
3.1	Introduction	27
3.2	Data Collection and Selection of Cases	29
3.3	SciHub Downloads Per-Country in the European Union	30
3.4	Downloads Per Academic Staff	33
3.5	Statistical Analysis of Data	34
3.5.1	Hypotheses and Method of Analysis	34
3.5.2	Findings	37
3.6	Discussion	40
4	Conclusion	47
	Bibliography	50
A	SciHub Data Analysis	59

List of Figures

3.1	SciHub downloads per month in EU 28 countries	31
3.2	SciHub downloads and library spending across the EU	38
3.3	SciHub downloads and published articles across the EU	39
A.1	Diagnostic plot for the regression of SciHub downloads per staff by library spending per staff	60
A.2	Diagnostic plot for the regression of SciHub downloads per staff by number of published papers per staff	62
A.3	Sci-Hub downloads and expenditure on tertiary students across the EU	64
A.4	Diagnostic plot for the regression of SciHub downloads per student by annual expenditure per tertiary-level student	65

List of Tables

3.1	Percent increase in total number of SciHub downloads across the EU between September 2015 and February 2016	32
A.1	Linear regression between number of SciHub downloads per academic staff and library spending per staff	59
A.2	Linear regression between number of SciHub downloads per academic staff and number of published papers per staff	61
A.3	Variance inflation factors for model with Downloads per staff as dependent variable.	61
A.4	Multivariate regression with SciHub downloads per academic staff as independent variable	63
A.5	Linear regression between number of SciHub downloads per student and annual expenditure per tertiary level student	64
A.6	Dataset Used in Analysis	66

Acknowledgements

My completion of this dissertation could not have been accomplished without the instructions of the teaching staff at the University of Glasgow, especially my supervisor Dr. Carl Knight. I offer my sincere appreciation for the learning opportunities provided at the University.

My studies would have not be possible without the financial support of Chevening Scholarships, the UK government's global scholarship programme, funded by the Foreign and Commonwealth Office (FCO) and partner organisations, to whom I express my gratitude.

Finally, I am grateful for the support and encouragement I received from my family and friends during my stay in Glasgow. My heartfelt thanks.

Chapter 1: Introduction

Literature sharing platforms that enable individuals to bypass a publisher's restriction to scholarly publications, collectively known as pirate libraries, have received considerable recent attention by both researchers and the public. The primary interest has centred around the amount of scholarly content held by pirate libraries, the motivations for their establishment, and the motivations for contributing towards them (Cabanac, 2015; Dunn, Coiera, and Mandl, 2014; Gardner, C. C. and Gardner, G. J., 2017). The general usage of pirate libraries has also gotten scholarly attention (Bohannon, 2016; Bodó, 2016b). However, very little work has been carried out to investigate why and how much pirate libraries are used by researchers (Cabanac, 2015). This dissertation addresses this research gap: it explores the usage of Sci-Hub.org (SciHub), currently one of the leading pirate libraries, by researchers¹ in the 28 member states of the European Union (EU). The investigation was made possible by the server logs data of SciHub that were made available to the public (Elbakyan and Bohannon, 2016).

The investigation, however, is focused on more than just usage. In providing justification for SciHub, Alexandra Elbakyan, the founder of SciHub, invoked Article 27 of the Universal Declaration of Human Rights (UDHR) (Torrent Freak, 2015c), which grants the right to 'freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits' (United Nations, 1948).

¹Throughout the text the terms researchers, scholars, academics, academic staff, teaching staff, and university staff are used interchangeably. They denote persons who are employed by universities as classroom teachers or academic staff (Eurostat, 2016).

This claim puts pirate libraries in a somewhat controversial situation. The legality of SciHub is being challenged in court (Elsevier, 2015; Torrent Freak, 2015a), but human rights struggles have always had a suspicious legal standing (Gordon, 2004) – at least until they have been won. The dissertation therefore investigates the assertion that the work of pirate libraries is addressing a human rights issue.

1.1 Aims, Objectives and Research Question

The aims of this dissertation are to evaluate the claim that limited access to scientific articles violates human rights, and to investigate to what extent does SciHub contribute to alleviating this violation. The specific objectives are:

1. To identify which human rights are undermined by limited access to literature and whose fulfilment would be improved by better access;
2. To present arguments that article sharing through pirate libraries is an activity that advances human rights;
3. To demonstrate that the needs for access of the scientific community are not sufficiently addressed;
4. To investigate usage of SciHub in the 28 European Union (EU) member states in order to evaluate its role in complementing access needs.

Specifically, this dissertation asks whether academic budgets influence the usage of SciHub by researchers across the European Union, and whether other factors such as researchers' publishing output, internet access, and piracy play a role.

The question is motivated by recent research regarding access to scientific literature across the EU. Tenopir, Volentine, and King (2012) found that academics from

six universities in the United Kingdom (UK) acquire around 30% of the needed articles and books via sources other than the university library or personal purchases. The UK has the second highest academic library spending per academic staff in the EU, and Italy, the third country in academic library spending per academic staff, spends 50% less than the UK (Online Computer Library Center, 2016). This suggests that even in the generally prosperous conditions found across the EU, the literature access gap could vary greatly between countries and universities. As a consequence, the need to complement existing library resources via non-conventional means might also be expected to vary.

Data from SciHub's servers show that the website is used everywhere in Europe (Bohannon, 2016). However, it is unknown how much SciHub's usage varies between countries and whether or not this variation is related to the investment in university libraries or other factors. Thus the research presented here focuses on SciHub's role in improving access to literature across the 28 EU member states and the scale of usage of this pirate library by researchers, neither of which has previously been addressed (Cabanac, 2015).

1.2 Summary of the Contents and Main Arguments of the Dissertation

The dissertation follows the stated objectives, presenting the interplay between restricted access to literature and human rights, before summarizing the insufficiency in access, and investigating SciHub's role in potentially bridging the access gap.

Chapter Two contains the literature review that addresses the first three objectives. The status of pirate libraries in the environment shaped by the current publishing model and intellectual property rights regime, and potential role of pirate libraries

in complementing shortcomings of the system cannot be fully addressed without an overview of the importance and consequences of access to literature or lack thereof. Of particular interest are the interrelated questions of whether or not limited access constitutes a violation of human rights, whether providing access by non-conventional means constitutes a human rights struggle, how important is access in the scholarly community, and how well does the current system meet the needs for access.

Chapter Three presents an analysis of data for SciHub downloads across the EU, and addresses the fourth objective that pertained to SciHub's usage and its potential to complement shortcomings in literature access in academia. To evaluate SciHub's role, its usage across the EU was contrasted to investment in access through academic libraries, and quantity of scientific outputs. Levels of internet access and protection of intellectual property rights were also considered as potential predictors of the number of articles downloaded from SciHub. The analyses focused on the EU because of the availability of standardized data sets that can be related to the SciHub data. In addition, although still variable, the political environment and prosperity levels across the EU can be considered stable relative to global variability. This removed factors with large expected effects on SciHub downloads, e.g., censorship, but allowed for a more narrow focus on the identified research gap regarding usage of pirate libraries in academia.

The findings showed that there was a general trend of increase in SciHub usage across the EU during the studied period. Investment in libraries was found to be poor predictor of the number of SciHub downloads, while publishing output positively affected SciHub usage. Internet access and software piracy have substantial individual effects on SciHub usage, but cancel each other's explanatory power when considered together.

Overall, the number and trend of increase in downloads from SciHub, combined with an apparent independence from investment, but dependence on research activ-

ity measured through publishing, suggested that SciHub contributes to bridging the literature-access gap caused by the current publishing model. As such, SciHub, and other similar services, might also help in the efforts to advance human rights undermined by restricted access to literature, most notably the right to enjoy the benefits of scientific progress and its applications (REBSPA), and the right to information (Byrne, 2007; Müller, 2010). However, the findings suggested that improving fulfilment of either right, and especially REBSPA, does not depend only on making more literature available, but requires additional efforts to improve research capacity and application of scientific discoveries.

Chapter Four concludes the thesis by synthesizing the presented material.

1.3 Pirate Libraries

Sci-Hub.org is a web site that allows users to search for and download scientific journal articles by entering the title or the digital object identifier (DOI) of an article. SciHub provides articles via proxy servers that, in turn, have access to repositories obtained through costly subscriptions to academic journals (Cabanac, 2015). Such articles are commonly called “pay-walled” in reference to the payment required to access them. The Library Genesis Project (LibGen) is a sister website to SciHub that provides hosting for articles. From a user’s standpoint, SciHub is the equivalent to a library service where one makes requests for specific books or journal articles, whereas LibGen is the equivalent of shelves of books and journals available in a library in the physical world. Pirate libraries is the general term used in this dissertation for LibGen, SciHub, and the other mentioned channels

Articles requested through SciHub are automatically copied for future availability to LibGen. The same is true for articles requested through other channels such as

the page Scholar (<https://www.reddit.com/r/scholar>) (Cabanac, 2015). It is also possible, although currently unknown, that requests made via the #icanhazpdf hashtag on Twitter.com are archived on LibGen.

In May 2016, SciHub provided access to over 50 million scientific articles (Bohannon, 2016). In March 2015, LibGen hosted 28 million documents, 95% of which were for educational purposes (Cabanac, 2015). The convergence of documents from these sources, suggests that by now, the volume of published material hosted on LibGen has probably greatly increased. To put these number into context, it is estimated that the total number of scientific articles ever published is around 50 million (Jinha, 2010). SciHub and LibGen, therefore, might provide universal access to a vast majority of the human scientific output to date, whose access is otherwise mostly restricted behind pay-walls. The rapid increase in the volume of data stored and shared in such non-conventional sharing platforms implies that these services might play non-negligible roles in the daily life of academia and the public. Elbakyan has stated that the goal of SciHub is to collect all scientific articles ever published, and to make them available for free (Torrent Freak, 2015c). The estimations above show that given time this ideal might not be unrealistic.

Chapter 2: Literature Review

2.1 Introduction

The literature review addresses the first three objectives of this dissertation, and provides theoretical context for the empirical analysis that follows.

The literature review is divided into three sections. The first section identifies the human rights that are undermined by limited access to scientific literature; it references scholarly work that argues for improving access to scientific literature as means to advance those human rights; and proposes the possibility for the advancement of human rights through providing access via pirate libraries. The section focuses on two rights, however given that human rights are considered interrelated, interdependent and indivisible (United Nations, 1993) the effects of advancing specific rights should not be viewed in isolation (Byrne, 2007, p. 115; Müller, 2010, p. 766).

The second section provides arguments that pirate libraries and articles sharing are a human rights struggle. This is done first by looking at pirate libraries as endeavours that come from the margins and challenge the established power relations which is a feature attributed to human rights struggles historically (Gordon, 2004, pp. 6-7). Thereafter, recent research is presented that investigates motivations of scholars who contribute to pirate libraries. The findings suggest that there are cases where the scholars' motivations echo the same human rights that have previously been identified to

be undermined by lack of access (Gardner, C. C. and Gardner, G. J., 2017, p. 11). Next, the critiques against pirate libraries are discussed. These include human rights issues for the authors, copyright infringement and financial issues for the publishers, and issues regarding metadata that libraries collect and process. Evaluation of these critiques arrived at the conclusion that most legal and ethical accusations aimed at pirate libraries are not justified.

The third section presents recent research about the importance and value of access for scholars, as well as the effects of access on educational results. This section demonstrates that even in rich countries such as the UK a substantial percent of the needs for academic literature are not met by university libraries (Tenopir, Volentine, and King, 2012, p. 135; Tenopir, Estelle, et al., 2015, p. 10). As the previous section of the literature review establishes, such lack of access undermines human rights. Therefore this section provides both background and context for the empirical investigation. The literature review concludes with a brief summary of the theoretical debate.

2.2 Human Rights That Can Be Advanced by Better Access to Scientific Literature

Elbakyan justified her work on SciHub through Article 27 of the UDHR (Torrent Freak, 2015c). However, she is not alone to link the issue of access to scientific literature to rights prescribed in human rights law. The case for improving access in order to achieve greater fulfilment of human rights in scholarly work been made directly with regards to two rights as defined in human rights documents: the right to enjoy the benefits of scientific progress and its applications (REBSPA), and the right to information.

Although the two rights are discussed separately below, in reality it would be difficult the draw a line between enjoying scientific progress and having scientific in-

formation on which that progress could be based. Moreover, The United Nations (UN) considers human rights to be interrelated, interdependent and indivisible (United Nations, 1993). There are arguments that fulfilment of REBSPA and the right to information influences the realisation of other human rights, including the rights to freedom of expression, to hold opinion, to an adequate standard of living, to education, to food, and to health (Byrne, 2007, p. 115; Müller, 2010, p. 766). Therefore, the approach taken here is for simplicity in characterizing the relationships between access limitations and the two main human rights affected by it, but with the full understanding that this is a unified subject.

There is scholarship that considers libraries as institutions that can advance many human rights. Samek (2007, pp. 23-25) proposed a list of human rights that includes almost all of the rights granted with the UDHR as issues that can be advanced by libraries. Jaeger, Taylor, and Gorham (2015) argued that the role of libraries cannot be boiled down to mere distribution of goods, but that libraries have societal functions such as education and services, and that are an equalising force in society and symbol of equality and justice (Jaeger, Taylor, and Gorham, 2015, p. 51). Therefore libraries provide variety of human rights services even though those services might not be labelled with that name (Jaeger, Taylor, and Gorham, 2015, p. 67). Both by declaration and in practice the services provided by either SciHub or LibGen overlap with the stated societal roles of conventional libraries. This strengthens the case that, like any other library, pirate libraries provide services that advance human rights via providing access to, perhaps, otherwise inaccessible documents.

2.2.1 The Right to Enjoy the Benefits of Scientific Progress

REBSPA is enumerated in two human rights documents. The UDHR in Article 27(1) grants the right 'freely to participate in the cultural life of the community, to enjoy the

arts and to share in scientific advancement and its benefits' (United Nations, 1948). The International Covenant on Economic, Social and Cultural Rights (ICESCR) in Article 15(b) grants the right to 'enjoy the benefits of scientific progress and its applications' (United Nations, 1966b).

Despite a broad consensus among states, REBSPA remained obscure long after the adoption of the UDHR and ICESCR (Müller, 2010, p. 766). The neglect of the right, at least from the side of the UN, ended in 2007 when the United Nations Educational, Scientific and Cultural Organisation (UNESCO) set up an initiative with the goal to clarify the normative content of REBSPA. The result of that initiative is the Venice Statement published in July 2009.

The key points from the Venice Statement relevant to this dissertation are that: the right is applicable to all fields of science (Müller, 2010, p. 770); that scientific progress can contribute towards the fulfilment of other human rights (Müller, 2010, p. 772); that the current intellectual property regime can be an obstacle to freedom of scientific research (Müller, 2010, p. 774); that states have a duty to promote development of skills for scientific research (Müller, 2010, p. 781), and that the scientific community, along with states and companies, has 'responsibility to ensure support for scientific inquiry and dissemination of scientific knowledge, and to actively pursue capacity building on a global scale, particularly in those countries which are relatively inactive in this regard' (UNESCO, 2009, p. 2).

Müller (2010, p. 782) proposed that the Venice Statement is an important move towards giving prominence to REBSPA and defining state obligations in regard to REBSPA. However, comprehensive fulfilment of REBSPA goes beyond scientific research: the ultimate goal should be both about developing useful application of science and affordable and accessible distribution of science (Müller, 2010, p. 781).

The Venice Statement advocates greater involvement of the scientific community in advancing REBSPA. There are, of course, many ways to imagine involvement of the scientific community in delivering affordable and accessible distribution of science. For example, it is possible that scientists set up an organisation similar to Médecins Sans Frontières (MSF) that will provide emergency help in given situations; or they might set up an organisation similar to Geekcorps, but instead of just providing assistance in development of computer infrastructure to developing countries, broaden the mission to scientific fields that are needed in a given context. Free online access to scientific research is another way to disseminate science, and evidence is presented below that pirate libraries could not have been made possible without the involvement of parts of the scientific community. Thus access granted through these services could help advance the under-developed REBSPA that is granted in human rights law.

2.2.2 The Right to Information

The right to information is enumerated in two human rights documents. The UDHR in Article 19 grants the right ‘to receive and impart information and ideas through any media and regardless of frontiers’ (United Nations, 1948). The International Covenant on Civil and Political Rights (ICCPR) in Article 19(2) grants the right to ‘to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice’ (United Nations, 1966a). In both documents the right to information is related to the right to hold opinions and the right to freedom of expression.

The idea that the right to information might be advanced through providing access to books comes from the International Federation of Library Associations and Institutions (IFLA). IFLA – a not-for-profit organisation whose membership includes associations of library professionals and individual libraries – can be considered as a

representative organisation of the librarianship profession. IFLA established the Committee on Free Access to Information and Freedom of Expression (FAIFE) in 1997. FAIFE was set up to promote the freedom of information as a necessity for all other freedoms, including education, development, self-expression, political action, and entertainment (Byrne, 2007, p. 115). The motivation for this decision was the wish to advance the librarianship profession together with advancing human rights (Byrne, 2007, p. 67).

Phenix and Peña McCook (2005, p. 24) made similar points when they argued that the library profession, with its rich history of alignment with human rights issues, should advance its commitment to human rights in the way doctors and lawyers within their professions have done through organisations such as MSF and Amnesty International. Rights granted under Article 19 of the UDHR are referenced again as the specific human rights that are advanced by librarians (Phenix and Peña McCook, 2005, p. 24).

FAIFE's mission establishes libraries as human rights institutions from the standpoint of the librarians, and puts the right to information on the agenda of the library profession. However, there are limitations to the efforts to advance the right to information. These limitations include public security, national interest, personal privacy, data protection, intellectual property rights, commercial confidentiality, and the protection of the vulnerable (Weermantry, 1997 cited in Byrne, 2007, p. 117). Other more practical limitations come from the limited physical shelf-space, digital storage capacity, and budget that libraries have (Byrne, 2007, p. 116). In the context of this dissertation, of primary interest are the limitations that intellectual property rights and their pricing model pose on access to scientific literature, and the ways in which pirate libraries can help overcome that limitation.

The scientific community is aware of the limitations posed by the intellectual property regime. Byrne (2007, p. 120) argued that scholarly publishing, and especially academic journals, have been reinvented by the publishing industry as a business asset for commercial exploitation, rather than a public good. The system benefits the biggest publishers, and is detrimental for academics and universities (Suber, 2012, p. 29). The pricing model of academic journals causes problems even for the richest libraries. According to data from 2008, Harvard subscribed to 98,900 serials, the best funded research library in India to 10,600, and some universities in sub-Saharan Africa to zero (Suber, 2012, p. 30). The words of Robert Darnton, the former director of Harvard Library, are probably the best to describe the sentiment about the system:

We faculty do the research, write the papers, referee papers by other researchers, serve on editorial boards, all of it for free [...] and then we buy back the results of our labour at outrageous prices. (Sample, 2012).

To add to this there is the issue of electronic access to journal articles that creates usage restrictions for scholars. These restrictions sometimes place bigger limits to the freedom to use materials than the limits that existed in the paper era (Suber, 2012, p. 34).

The pricing and licensing models together confound the access problem of libraries, both in terms of obtaining access and in providing it to the public. One of the solutions that is offered to the dysfunctional system of academic publishing is 'open access' (Suber, 2012, p. 29). Open access is defined as literature that is 'digital, online, free of charge, and free of most copyright and licensing restrictions' (Suber, 2012, p. 4). Big repositories of such literature might solve the access problem for many university libraries.

However, it is questionable to what extent open access currently solves the access problem. Bodó (2015, p. 101) asserted that it is a slow-moving initiative with limited

practical effectiveness. Reports from 2015 suggest that the proportion of articles published as open access is 12% (Ware and Mabe, 2015, p. 11). Open access, thus, solves a fraction of the access problem while the vast majority of scientific literature remains inaccessible without paid subscriptions.

To summarise, libraries and librarians see themselves as promoters of human rights. Central to their mission is the right to freedom of information as granted under Article 19 of UDHR and ICCPR. With regards to scientific information, one of the main limitations to the right of information is the intellectual property regime that manifests itself through the pricing and licensing model of major publishers. Initiatives such as open access try to solve the pricing and licensing problems, arguably with limited success.

It is in this context that SciHub, and other services for obtaining free scientific literature, complement the libraries in their mission to advance human rights. Access granted through these services can help advance the right to information whose fulfilment is hindered by the intellectual property regime.

2.3 Article Sharing as Human Rights Activism

The Venice Statement and FAIFE recognise the current intellectual property regime as an obstacle to fulfilling the right to information and REBSPA. The previous section proposed that SciHub and other services that enable free sharing of scientific articles can help overcome this obstacle. Bohannon (2016) showed that SciHub is used across the globe suggesting that the work of pirate libraries complements efforts of other initiatives for providing better access and creates possibilities for greater fulfilment of REBSPA and the right to information. The following section explores in more detail whether the work of pirate libraries can be considered as a struggle for human rights.

2.3.1 Pirate Libraries as a Human Rights Struggle

Pirate libraries were originally set up as means to overcome Soviet-era censorship (Bodó, 2015, p. 99). The ICCPR was in force in the Soviet Union since March 1976 (United Nations, 1966a). Therefore, the early pirate libraries, the copying of physical books, and black market exchange of printed materials in the Soviet Union can be viewed as activism to exercise rights granted under the ICCPR in an environment where party censorship determined what could be read (Bodó, 2016c).

Most of the digital pirate libraries were and are located in Russia: from Gigapedia, to Aleph, to LibGen (Bodó, 2016b). The digital pirate libraries provide a solution to the price and access problem of academic journals and books by providing universally unrestricted access. (Bodó, 2015, p. 101). Bodó (2015, p. 101) argued, that in the context of the East–West division in Europe, the goal of pirate libraries is to provide Eastern Europe the same level of access available in the Western World, and to serve as agents of modernisation.

The spread of the internet means that pirate libraries enable universal access across the globe and this shortcut to universal access challenges the established intellectual property regime (Bodó, 2015, p. 101). This radical approach can be viewed as an opposition to the notion that Western countries use the intellectual property regime to maintain global inequality (Karaganis, 2011 cited in Bodó, 2015, p. 101), specifically with regards to publishing and transfer of knowledge (Helfer and Austin, 2011, p. 340).

If the intellectual property regime is used as means to maintain global inequality, then it can be argued that pirate libraries represent a human rights struggle from the margins against the power structures (Gordon, 2004, p. 6). Gordon (2004, p. 7) argued that throughout history human rights emerged from oppression and exclusion. Lack of access as a consequence of the current pricing and licensing models results in

exclusion of large number of people from humanity's pool of knowledge. Helfer and Austin (2011, p. 237) argue that REBSPA cannot be reduced from the right to benefit from the effects of science to the right *to wait to benefit* from the effects of science. The same can be said regarding the right to information that is a prerequisite for the right to freedom of expression and the right to hold opinions (Byrne, 2007, p. 115). Thus, pirate libraries offer an avenue for inclusion despite the danger of prosecution by the established powers in trade relations, and the improved access to literature can be considered an outcome of the struggle for human rights. Improving access is in line with the recommendations of the the Venice Statement and the FIAFE project.

2.3.2 Motivations for Contribution to Pirate Libraries

The Venice Statement called for intensified involvement of the scientific community in the efforts to fulfil REBSPA. It appears that through her actions, Elbakyan, who is a computer scientist and a member of the research community, answered this call. However, the ever-expanding volume of articles that are available on SciHub was not compiled by one person.

Focusing on one scientific field, Dunn, Coiera, and Mandl (2014) explored the possibility of 'biblioleak' – a breach in the repositories for biomedical articles that are at risk just as any other content generating industry. Dunn, Coiera, and Mandl (2014) argued that a younger generation of researchers for whom free access is a desired norm might have the motivation for releasing a biblioleak. This belief is founded in the culture of scientific communities that for a long time has nurtured the idea that 'information should be free' (Dunn, Coiera, and Mandl, 2014). A second reason for a biblioleak, at least as far as biomedical research is concerned, might be found in the low rates of self-archiving by researchers even in cases where doing so will not constitute copyright infringement (Dunn, Coiera, and Mandl, 2014). Self-archiving

is the practice of researchers making articles available on publicly available personal websites. High rates of self-archiving would make a leak less useful.

It seems that biblioleaks indeed have happened. Cabanac (2015, p. 878) found that LibGen has benefited from massive contributions to its database at a rate of over 100,000 articles per day in a 13 days period. There is no way to verify the origin of those files, but the most likely scenario is that they came from a content database available at a university library or similar institution that has broad access. This suggest involvement of person(s) from the scientific community. The daily contributions to LibGen have a median value of 2,720 articles (Cabanac, 2015, p. 878). The most likely scenario again is that a sizeable portion of the daily additions come from members of the scientific community that have access to databases, or at least from academics who gained access to articles through friends or colleges. The motivations for these large amounts of contributions to LibGen can not be established, but it is unlikely that they happened unintentionally, so it is possible that the contributors might have had political motivations.

This is not to say that all of the academics who are involved in the sharing of articles do that to advance REBSPA or the right to information. A yet to be published research of scholars' practice in downloading and uploading journal articles, found that most respondents cite reciprocity when explaining their motivations for uploading to a sharing platform (Gardner, C. C. and Gardner, G. J., 2017, p. 11). Reciprocity in this context means providing content back to the broader community from which a previous download has occurred (Cenite et al., 2009, p. 210).

However, about a third of the 104 scholars that responded in the study provided an answer that was coded as ideology (Gardner, C. C. and Gardner, G. J., 2017, p. 11). Gardner, C. C. and Gardner, G. J. (2017, p. 3) reported that moral issues and ideology are more prevalent among respondents who upload articles, when compared to those

who download articles. Gardner, C. C. and Gardner, G. J. (2017, p. 3) did not offer any explanation as to what that ideology might be, but they do refer to Aaron Swartz's Guerrilla Open Access manifesto in describing the answers of some of the scholars that participated in the study. Aaron Swartz's Manifesto calls to 'download scientific journals and upload them to file sharing networks' (Swartz, 2008) in order to provide access to those who do not have it. Therefore, it is possible that some of the scholars share articles in order to challenge the current publishing model that is one of the main reasons for limited access.

A study by Tenopir, Hughes, et al. (2014) among scholars in the UK and the United States of America (USA) also reported on motives for sharing journal articles among university researchers and students. The motives include furthering scientific discovery and fulfilling need for information for those who have no access, as well as practical concerns such as sharing out of convenience (Tenopir, Hughes, et al., 2014, p. 126). Similar findings are reported by Tenopir, Estelle, et al. (2015, p. 5) this time from a study that, in addition to the UK and USA, included scholars from Canada, India, China, and Italy. Motivation such as furthering scientific discovery and fulfilling needs for information echo the goals set in the Venice Statement and FAIFE.

Both studies found email to be the most used channel for sharing, although Twitter and other social media websites play a role as well (Tenopir, Hughes, et al., 2014, p. 123, Tenopir, Estelle, et al., 2015, p. 4). While these studies do not suggest intentional usage of SciHub, the pirate library is set up in a way that enables unintended contributions to the pool of accessible articles. Cabanac (2015) showed that article requests that occur on the website Reddit are eventually archived on LibGen. It is possible that the same happens with articles that are made available elsewhere on the internet. Thus, even when someone answers a request just to be helpful, the shared article might be copied into the expanding LibGen catalogue. The same happens when a user searches for an article for her own needs: once SciHub provides the article to the user, it auto-

matically uploads a copy of it to LibGen ensuring future availability (Cabanac, 2015, p. 881). The overall effect of the article exchange is a creation of a substantial library of scientific literature available to everyone.

In summary, the work of pirate libraries can be considered a struggle for human rights, and many participating individuals have motivations that echo REBSPA and the right to information as their motivation for contributions to pirate libraries. Furthermore, many of the scholars who share articles without directly participating in the work of pirate libraries, might do so with a similar rationale related to human rights.

2.3.3 Critiques Against Pirate Libraries

There are three arguments that critique pirate libraries. First there is the question of whether pirate libraries infringe on the ‘right to protection of moral and material interests resulting from any scientific, literary or artistic production’ that is granted to authors in Article 27(1) of the UDHR, and Article 15(c) of the ICESCR (United Nations, 1948; United Nations, 1966b). These provision in the UDHR and the ICESCR seem to oppose REBSPA opening the possibility for an ‘internal conflict’ – a situation where the same legal document contains provisions that seem to be conflicting each other (Foster, 2008, p. 303)

However, as far as scientific resources are concerned, pirate libraries do not remove authors’ names or other information from articles, nor they endanger the integrity of the work. To do so would endanger the credibility of the scientific work and defeat the purpose of having a pirate library-enabled access in library-poor regions. Thus, the moral rights of authors that are granted as a human right under UDHR and the ICESCR are not in question.

It also can be argued that pirate libraries do not endanger material rights of authors that are granted as a human right under UDHR and ICESCR. Authors usually do

not get paid for writing nor they receive royalties from publishers (Mars and Medak, 2015 cited in Bodó, 2016a, Suber, 2012, p. 37), regardless of whether an article is available in a pirated format or not. Therefore, this conflict does not seem to exist: authors' moral and material interest are not endangered by the work of pirate libraries.

Second, and as evident from the law suit against SciHub, there is a claim that pirate libraries represent a copyright infringement scheme that costs publishers money. Publishers' rights are guaranteed through copyright law domestically, and trade agreements such as the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) internationally (Foster, 2008, p. 299). This creates another potentially conflicting situation. On one hand TRIPS and copyright law place restrictions on distribution of protected good and services. On the other hand, human rights law grants REBSPA and the right to information.

The situation in which human rights law seems to conflict other legal documents is called external conflict (Foster, 2008, p. 304). In resolving this issue Foster (2008, p. 305) discusses the World Trade Organisation (WTO) three part balancing test that aims to evaluate the conflict. In Foster (2008, p. 305) the analysis of the test is regarding the right to education, but by analogy a similar reasoning can be applied to REBSPA and the right to information. The three-part test includes evaluating:

- 1) the importance of interests or values that the challenged measure is intended to protect; 2) the extent to which the challenged measure contributes to the realisation of the end pursued by that measure; and 3) the trade impact of the challenged measure (Foster, 2008, p. 305).

Foster (2008, p. 305) argued that the importance of education is high; that free or cheap access to printed educational materials is an important contribution for better education in depressed economies; and that the trade impact in a depressed economy with no ability to pay is negligible. This would tip the balance towards proving access rather than protection.

In transferring this analysis to REBSPA and the right to information, it can be argued that the importance of benefits of science and information is also high. However, the test gets more complicated with the trade impact condition given that the pirate libraries exist online and thus they are not limited to economies with no ability to pay.

If access provided by pirate libraries causes conventional libraries to cancel subscriptions, then publishers might indeed lose money in richer countries. However, the frequency and reasons for any cancelled subscriptions are currently unknown, and criticism for such cancellations might be leveraged at both pirate libraries McNutt (2016) and expensive subscription fees (Sample, 2012). In poorer countries, where libraries have no ability to pay the asking subscription prices, there is no trade impact because there are no lost transactions. In this case, as Foster (2008, p. 306) argued, there exists a false conflict, and the WTO test would tip the balance in favour of providing access again.

To summarise, this particular claim against public libraries seems to have some substance. There is a possibility that in some markets earnings are lost. However, this can not be generalised across the globe and maybe a case by case investigation could provide evidence for discontinuation of subscriptions because of pirate library access.

The third argument against pirate libraries is that there is more at stake than loss of profits. McNutt (2016) argued that when people use pirate libraries, information about article usage and download statistics that are useful measures of the impact of the work are lost. This seems to be a weak claim having in mind present day technology. Tracking of article and similar issues are simple technical features that in time can be implemented without complication to SciHub, or any other pirate library. SciHub already has the Digital Object Identifier (DOI) of each article it holds (Elbakyan and Bohannon, 2016). Recording meta-data about article usage that can be accessible to authors is a trivial task in the endeavour for providing access. And even without having

these features there is a possibility that authors might benefit. Dunn, Coiera, and Mandl (2014) proposed that articles leaked to pirate libraries have no negative consequences for researchers, and in fact result in improved visibility of the work for the said authors.

2.4 The Value and Benefits of Access

The arguments above showed that there are human rights that are undermined by limited access, that there are people who have those rights in mind when contributing to pirate libraries, and that the critique leveraged against pirate libraries is generally not justified. The following section examines the literature needs of the scientific community.

Journal articles are an important resource for academics both for teaching and research. In a study regarding the value of libraries to academics in the UK, Volentine and Tenopir (2013, p. 429) found that many considered journal articles to be important for research and teaching, and described them with words such as ‘critical’, ‘essential’, or ‘vital’. At the same time, many respondents raised issues with access to journal articles through the library services of their universities (Volentine and Tenopir, 2013, p. 429).

Another study by Tenopir, Volentine, and King (2012, p. 131) found that academics in six UK universities base about half their reading in articles: of the average 39 scholarly readings per month 22 are journal articles. An interesting finding is that in accessing scholarly materials, academics use other sources, i.e., not library or personal purchase, in about 30% of cases of both articles and books that they read (Tenopir, Volentine, and King, 2012, p. 135). The authors did not pursue what ‘other sources’ means, but it is not unreasonable to assume that pirate libraries might play a role in providing needed materials. In any case, this is a clear indicator of the insufficiency of university libraries even in rich countries such as the UK.

More recent research of article usage in the UK and USA (Tenopir, Estelle, et al., 2015) confirmed both the importance of university library access for obtaining journal articles, and the limitations that libraries have. About 2/3 of articles needed for teaching and research come from the library to which academics have access (Tenopir, Estelle, et al., 2015, p. 10). Furthermore, Tenopir, Estelle, et al. (2015, p. 10) estimated that for every recorded download, an article could be shared for an additional ten times. The download:share ratio might be an even better measure of the importance of journal articles in the work of academics than the officially recorded downloads.

Access to instruction materials, though not at the university level, has been established to be of big importance for achieving good results. The impact on test scores per dollar spend on educational inputs has been measured in Brazil in the 1980s, and in India in 1990s. In both cases increases in access to materials have resulted with significantly higher increases in test scores when compared to teacher's salary, teacher's education or available facilities in school (World Development Report, 2004 cited in Foster, 2008, p. 288).

However, despite these findings, access to literature in libraries, and in turn to scholars, around the world has remained limited. A study regarding access to knowledge and copyright in eight African states suggested that developing countries face challenges in access to academic literature both at institutional and personal level. Poor resources in the libraries are coupled by inability of library users to buy literature themselves because of high prices set by publishers (Armstrong et al., 2010, p. 335). These findings are reported despite the possibility that developing countries can receive free or discounted access to some scientific databases (Research4Life, 2016), suggesting a persisting problem of access to literature.

With access playing such an important role for scholars and its relative scarcity at different institutions and across states, alternative ways to access journal articles

might complement the existing resources to conduct scientific research and expand the distribution of scientific benefits. Pirate libraries, such as SciHub, might provide what is missing for many scholars around the world.

2.5 Conclusion

The literature review focused on the first three objectives of the dissertation. It identified the human rights are undermined by limited access to literature and whose fulfilment would benefit from improved access to information; it offered arguments that article sharing through pirate libraries is an activity that advances human rights; and it demonstrated that the needs for access of the scientific community are not sufficiently addressed.

The right to enjoy the benefits of scientific progress and its applications (REB-SPA) and the right to information are the human rights, defined in international human rights documents, are directly and most severely affected by limited access to literature.

Efforts to improve access to literature have emerged as mechanisms of human rights struggle, and are claimed as such, by both flagship librarianship organizations and scholars that have developed and maintain pirate libraries. Pirate libraries originated as human rights struggles in oppressive societies, but thanks to the internet have expanded their scope from local or regional devices meant to oppose censorship to larger efforts aiming to improve access to literature globally. Previous research has found that sharing articles is often politically motivated, and that the human rights identified above are among the primary reasons for direct contributions to pirate libraries. Furthermore, even in the countries that spend the most on libraries, the needs for scientific literature remain insufficiently addressed. At the same time scholars consider scientific articles as vital and essential for their work. This strengthens the position

of pirate libraries as providers of much needed access, that, as a consequence, helps alleviate human rights violations in an age when access to literature is of paramount importance. Taken together the efforts of running and contributing to pirate libraries might be likened to historical human rights struggles that challenged established power relations.

Chapter 3: Analysis of SciHub Data

3.1 Introduction

The empirical analysis of SciHub usage aims to address the fourth objective of the dissertation: to investigate usage of SciHub by researchers in the 28 EU member states in order to evaluate SciHub's role in addressing access needs. This is approached by investigating the relationships between number of downloads, on one hand, and spending on libraries, number of published articles, and the availability of the service accounted through access to internet and prevalence of piracy, on the other. Beside explaining the usage of SciHub, the results can offer insight into the potential of pirate libraries for bridging the access gap and making the discussed human rights violations less severe.

The motivation for this approach is found in recent research. Tenopir, Volentine, and King (2012, p. 135) found that scholars in the UK turn to sources outside university libraries or personal purchases for roughly one third of the needed literature. The author is not aware of similar research conducted in other EU member states. However, the UK data show that the second highest investment in library spending per academic staff in the EU is sufficient for about 70% of the needed literature. Only Finland spends more per academic staff than the UK, and third in line is Italy with 50% of UK's spending (Online Computer Library Center, 2016). If one assumes similar needs

in terms of number of books and articles, as well as equal subscription fees across the EU, then the fraction of inaccessible literature should be higher in countries that spend less on academic libraries.

The statistical analysis begins with the relationship between number of downloads from SciHub and academic library spending. Academic library spending is used as an indicator of the available access in libraries: more spending should lead to better access. As a consequence the need for alternative sources should be greater for academics in EU countries that spend less on academic libraries. In those countries pirate libraries might have a more prominent role as an alternative source for access, and as a result, downloads from SciHub might be more frequent.

However, usage of pirate libraries might be influenced by other factors. First, it is probably safe to assume that more productive researchers need more articles, and in a climate of restricted access, might end up using SciHub more. Also, downloads might be determined by other factors not related to the research community, but to the country's legal environment. To account for these possibilities, the analysis takes into consideration two additional factors: percent of internet access as the measure of the availability of SciHub's website, and percentage of software piracy as a measure of the protection of intellectual property rights across the EU.

The EU was the primary focus because of the availability of standardized data that can be related to SciHub usage. Moreover, the political environment across the EU is relatively stable which removed potential factors with large expected effects that could have affected the analyses (e.g., political or religious censorship). Finally, although prosperity levels across the EU are comparable, differences in the degree of development of the education and research systems still exist, especially along the axis of before vs. after 2001 membership (Bodó, 2015). It was therefore of interest to

evaluate whether pirate library usage might be potentially important in the effort of overcoming these differences.

This chapter first presents the collected data and summarises SciHub downloads per EU country and per academic staff. Descriptions of the hypotheses and methods used to test them follow, before the results and a discussion of the main findings.

3.2 Data Collection and Selection of Cases

SciHub data for number of downloads per country between 1.9.2015 and 29.2.2016 were taken from the SciHub data set (Elbakyan and Bohannon, 2016). The analysis, was restricted to EU member states because Eurostat (Directorate-General of the European Commission) provides standardized data sets for EU members for a number of variables of potential relevance to the study.

Data for number of tertiary education staff per country is for year 2014 (Eurostat, 2016). Teaching staff refers to persons who are employed by universities as classroom teachers and academic staff (Eurostat, 2016).

Data for number of published articles per-country were obtained from the publicly-available Scimago Journal and Country Rank (Scimago Journal & Country Rank, 2016), which, in turn, is derived from Scopus – Elsevier’s abstract and citation database. These data were for the year 2015.

Data for spending on academic libraries were taken from the data set on Global Library Statistics provided by Online Computer Library Center (Online Computer Library Center, 2016). These data were used to account for the regular access to journals that is available to academic staff. These data were not from a single year: they ranged from 1995 to 2014 and, often, the entry for the number of academic libraries per country is not for the same period as the entry for library spending. However, provided no

major reforms of the educational system took place in EU member states, the number of universities, and therefore university libraries, is not expected to vary enough such that it makes these data incompatible. Also, academic library spending data were unavailable for Cyprus. Cases where there are missing data were omitted from all analyses.

Data for percent of internet access for the year 2014 were taken from The World Bank (The World Bank, 2014). These data were used as an indicator of availability of SciHub. However, internet access alone is not the best measure of availability of web resources. For example, the UK has 92% internet access, however, SciHub's website in the UK is blocked by the internet providers. As an alternative way to account for the intellectual property rights environment in each country, data for rate of software piracy were taken from The Business Software Alliance study for 2016 (Business Software Alliance, 2016).

3.3 SciHub Downloads Per-Country in the European Union

Figure 3.1 on page 31 shows the download trend for the EU as a whole. Downloads from SciHub across all 28 EU member states had grown between September 2015 and February 2016. The steep fall in November 2015 is most likely due to the SciHub domain being suspended making the website inaccessible for most users (Torrent Freak, 2015b).

Nevertheless, by the end of the six month period for which data were available (February 2016), the combined total EU downloads and downloads in each country were at a higher level than in September 2015 (Figure 3.1; Table 3.1 on page 32). Table 3.1 shows the percent increase per country in the EU. The column EU Code

denotes which countries were members of the EU before the eastern expansion of the bloc in 2001, and which countries joined after 2001. This division is important in the context of the need of Eastern European countries catch up with their western neighbours (Bodó, 2015, p. 101), although the eastern expansion of the EU includes countries such as Malta which did not belong to the Soviet bloc. In any case, Table 3.1 shows that the increase in downloads cuts across this division. For example the countries with greatest increase were Estonia – a former Soviet republic, and Finland. Also, countries such as the UK, the highest spender on academic libraries in the EU, and Romania, one of the most modest spenders, have about the same percent increase in SciHub downloads.

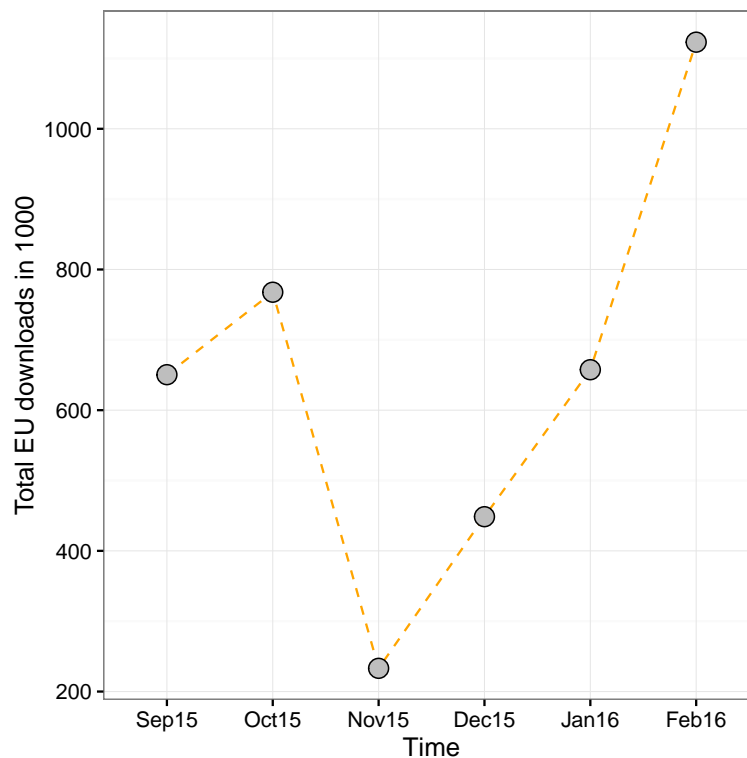


Figure 3.1: SciHub downloads per month in EU 28 countries. The sharp decline between October and November 2015 was due to Sci-Hub.org being suspended. In the four subsequent months with available data, SciHub usage across the EU rose by 382% in Source of data Elbakyan and Bohannon (2016).

Table 3.1: Percent increase in total number of SciHub downloads across the EU between September 2015 and February 2016. Source of data Elbakyan and Bohannon (2016).

EU Group ^a	Country	Total downloads ^b	Percent increase ^c
EU13	Estonia	8072	302.6
EU15	Finland	20402	279.6
EU13	Romania	108673	238.0
EU15	United Kingdom	218631	232.4
EU13	Malta	6687	223.7
EU13	Slovakia	26930	199.5
EU15	Sweden	29373	169.7
EU15	Luxembourg	11958	164.5
EU15	Denmark	10864	161.4
EU13	Czech Republic	63983	154.9
EU15	Ireland	74381	139.7
EU13	Poland	137888	138.6
EU13	Slovenia	40809	120.5
EU15	Belgium	129615	118.8
EU15	Austria	61566	115.0
EU15	Spain	426412	107.1
EU13	Lithuania	44260	103.9
EU15	Italy	295015	99.2
EU13	Croatia	75514	83.2
EU13	Hungary	119730	80.5
EU15	Greece	285452	75.5
EU15	Portugal	434721	47.2
EU13	Bulgaria	67289	45.7
EU13	Cyprus	4484	40.1
EU15	France	511145	39.1
EU13	Latvia	35693	38.5
EU15	Netherlands	168876	38.2
EU15	Germany	462011	1.4

^aEU group refers to the date of joining the union, before or after the eastern expansion in 2001.

^bSum of downloads over the six month period for which data were available.

^cPercent increase of SciHub downloads between September 2015 and February 2016.

3.4 Downloads Per Academic Staff

The total number of downloads might be a good initial indicator of SciHub's usage. However, the numbers have to be put into context in order to evaluate the extent of the access gap in academia, and how much SciHub contributes to bridging it. A study conducted in six UK universities found that academic staff read on average 22 articles per-month (Tenopir, Volentine, and King, 2012, p. 131) adding up to an average of 132 articles over a six month period. A total of 218,631 downloads from SciHub were made from the UK during the six month period for which data were available (Elbakyan and Bohannon, 2016). This means the 151,566 academic staff in the UK (Eurostat, 2016) downloaded, on average, 1.4 journal articles per staff in the given period. Assuming that the average six month reading is similar across UK universities, these 1.4 downloads per staff seem relatively modest.

However, given the possibility that an article is being re-shared up to ten times after downloading (Tenopir, Estelle, et al., 2015, p. 10), the total effect of SciHub downloads in the UK might be greater. Also, it should be noted that SciHub's website remains blocked by major internet providers in the UK. This means that researchers need greater effort or technical proficiency to access the site. It is also possible that academics in the UK turn to other non-conventional article sharing channels, such as the Twitter.com hashtag #icanhazpdf or Reddit.com/r/scholar page, more often.

The highest number of downloads per staff were recorded in Greece, Luxembourg, and Portugal with 18.75, 14.97, and 12.97 downloads, respectively. To the best of the author's knowledge, data about the average reading by academic staff in these countries are not available. However, if the average of 132 journal articles in a six month period is similar across the EU, then SciHub might play a more important role for the academic communities in these countries. This might be especially true for

Greece and Portugal where both the number of total downloads and downloads per staff are high.

3.5 Statistical Analysis of Data

3.5.1 Hypotheses and Method of Analysis

The hypotheses were derived from the relationships discussed above and reflect the expectations that SciHub usage should depend on the quantity of access provided by academic libraries as well as the frequency of publishing by academic staff.

1. There is a significant relationship between library spending and downloads from SciHub;
2. There is a significant relationship between publishing scientific articles and downloads from SciHub;

The data were analysed with bivariate and multivariate regressions and the major aim was to test how aspects of spending (on libraries or students) or publishing explain SciHub downloads. Prior to analyses, the data were converted to the same units, e.g., SciHub downloads per teaching staff vs. library spending per teaching staff. In general form the hypotheses can be summarised by the following statements:

$$H_0: \beta_i = 0$$

$$H_a: \beta_i \neq 0, \text{ where } \beta_i \text{ stands for the slope coefficient(s).}$$

Each country of the EU28 was a case. Countries with missing data for particular variables were removed from the analysis. Significance was tested at $\alpha = .05$, with p-values smaller than α taken as support for rejecting the null hypothesis.

The presence of outliers was tested using the standardized residuals of the regressions correcting for multiple comparisons using the Bonferroni correction (Field, 2009, p. 215). The analysis was carried out using the R statistics software (R Development Core Team, 2016).

The variables used in the first statistical test were number of downloads per teaching staff (calculated as number of downloads divided by number of staff) and academic library spending per teaching staff (calculated as academic library spending divided by number of staff).

In the first test, a result that would fail to reject the null hypotheses would indicate, simply, that SciHub usage is independent of the amount spent on academic libraries, whereas rejecting the null hypothesis, would suggest a relationship. There are two levels at which this test is relevant to the broader question of the importance of SciHub in complementing access needs. First, a significant negative relationship, in which the number of SciHub downloads increases while library spending decreases, would support the expectation that academics in countries that spend less on libraries are more likely to access SciHub for otherwise inaccessible literature. Second, if the two variables were not significantly related, then one can make the inference that SciHub's role in academia is of similar importance regardless of the amount of money spent on libraries across the EU. In the latter case, judging the importance of SciHub becomes a matter of quantity of downloads – lots of download across various countries would support an important role, and *vice versa*. However, just how many downloads are enough to label the service important with respect to alleviating access needs is difficult to determine.

The second statistical test investigated the relationship between number of downloads per staff and number of published articles per staff. In this test, the number of

published articles per staff was calculated as number of published articles divided by number of staff in each country.

In the second test, a significant positive relationship would indicate that SciHub usage grows as the productivity of the academic community in a country grows. If, on the other hand, there is a negative relationship, then one can make the inference that researchers in countries that publish less, are more likely to download from SciHub, perhaps as a consequence of the lower overall support for research and by extension more severe lack of access to literature. In the case of no relationship between SciHub downloads and publishing, the interpretation would again depend on the quantity of downloads. A point can be made, in this case again, that SciHub is of similar importance regardless of the scientific output of a country. Such a result might indicate more subtle effects of the need for access. For instance, highly productive researchers in well funded institutions might nonetheless obtain access to more obscure literature through SciHub. On the other hand, scholars in countries with weaker research communities might frequent SciHub for general access to articles from obscure or mainstream journals. In the end, SciHub usage might be comparable despite substantial differences in publishing output.

A multivariate regression analysis was conducted to assess the effects of predictors relative each other and account for additional factors that might influence SciHub usage. Two other independent variables were added to the model: rate of software piracy and percent of internet access. The reason for the two new variables is the possibility that level of internet access might influence the access to SciHub as it is basically a web site. However, given high protection of intellectual property rights, websites that provide access to copyrighted content might be blocked (as is the case in the UK). Software piracy is thus also added to the model.

3.5.2 Findings

A visual investigation of the relationship between downloads per teaching staff and academic library spending per teaching staff showed a weak negative relationship: as library spending went up, downloads went down (Figure 3.2 on page 38). The relationship however was not significant. It was found that academic library spending per staff does not explain a significant amount of the variance of SciHub downloads per staff ($F(1, 25) = 1.494, p > .05, R^2 = .06$). Thus the null cannot be rejected ($\beta_1 = 0$). The standardized residuals test detected Greece as an outlier. However, removing Greece from the model did not result in a significant change ($F(1, 24) = 2.079, p > .05, R^2 = .08$) (Table A.1 on page 59).

In addition to library spending per staff, preliminary analyses also tested expenditure on tertiary education (undergraduate and postgraduate studies) as a possible predictor of SciHub downloads. This factor was similar to spending on libraries, in that it broadly accounted for investment in educational and research resources. However, the relationship between spending on tertiary education and SciHub downloads was not significant and did not offer insight different from expenditure on academic libraries. Therefore it was omitted from the main text, but the results are presented in the appendix (Figure A.3 on page 64; Table A.5 on page 64).

A visual investigation of the relationship between downloads per teaching staff and published articles per teaching staff showed a moderate positive relationship: as published articles went up, downloads went up as well (Figure 3.3 on page 39). The relationship was significant ($F(1, 26) = 5.928, p < .05, R^2 = .186$). It was found that published articles per staff explain about 18% of the variance of SciHub downloads per staff. Thus the null hypothesis can be rejected ($\beta_1 \neq 0$). (Table A.2 on page 61).

Given the residuals, Greece was again an outlier, but as before its removal had a minor effect on the model, which remained significant explaining a similar amount of

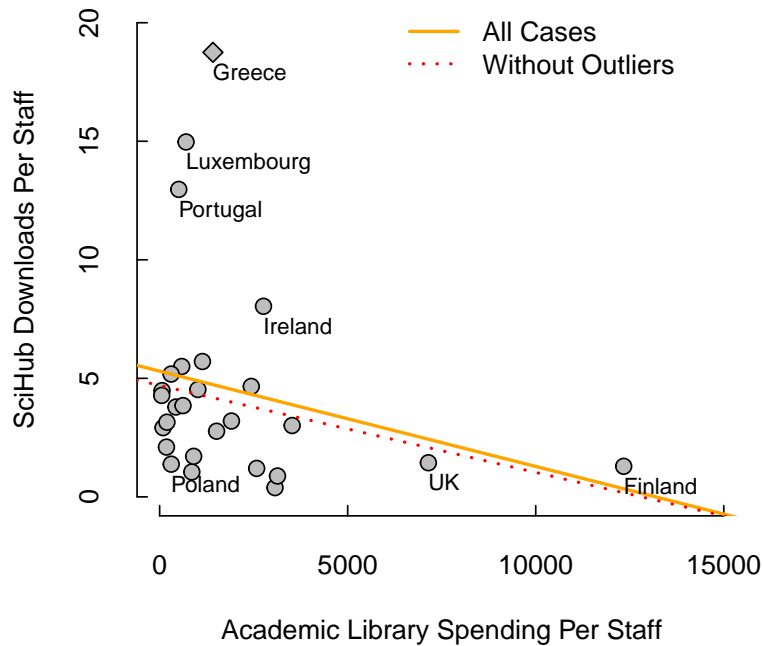


Figure 3.2: Linear regression of SciHub downloads per academic staff and library spending per academic staff across the EU. All cases (full line) and without (dashed line) the outlier country Greece.

variance in SciHub downloads ($F(1, 25) = 5.781, p < .05, R^2 = .188$). In both models published articles per staff statistically significantly predicted the value of downloads per staff ($\beta_1 = 4.346, t = 2.435, p < .05$) and ($\beta_1 = 3.399, t = 2.404, p < .05$), respectively. The effect of the predictor was that at the level of individual academic staff, a unit increase in number of publications was accompanied by about four downloads from SciHub (Table A.2)

To fit the multivariate regression to the data, a statistical analysis was conducted in order to find the best subset of the variables for predicting SciHub downloads per staff (Table A.3 on page 61). The results showed that the model with of published articles per staff, library spending per staff, and percent of internet access (Model 3 in Table A.3) explained a significant amount of the variance in SciHub downloads – around 48% ($F(3, 23) = 8.983, p < .05, R^2 = .540, R^2_{Adjusted} = .479$), and provided the best fit when compared to the other models. Furthermore, two of the coefficients

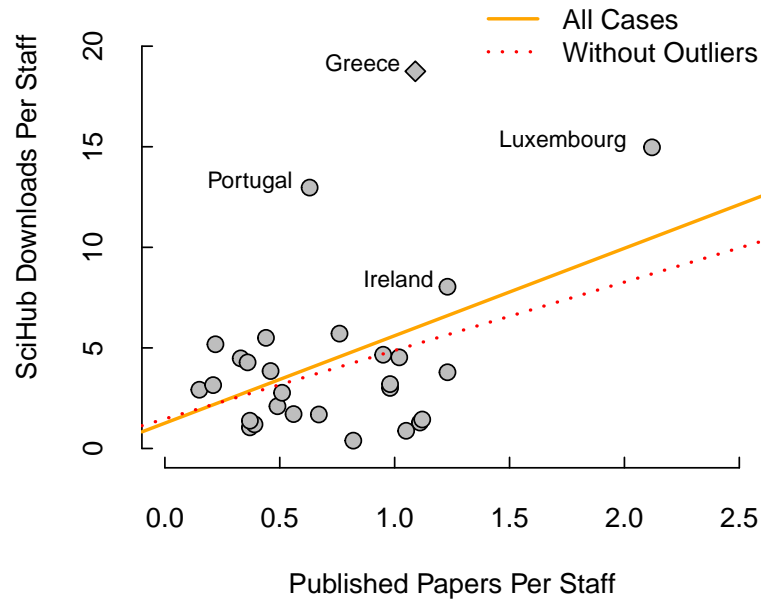


Figure 3.3: Linear regression of SciHub downloads per academic staff and number of published articles per academic staff across the EU. All cases (full line) and without (dashed line) the outlier country Greece.

were also statistically significantly different from zero: published articles per staff and percent of internet access statistically significantly predicted the value of downloads per staff ($\beta_1 = 8.082, t = 4.809, p < .05$) and ($\beta_3 = -18.890, t = -2.933, p < .05$), respectively).

Further investigation of the subsets of independent variables revealed that when rate of software piracy was kept in the model, both the effect size and significance of internet access changed ($\beta_3 = -14.459, t = -1.363, p > .05$) (Model 1 in Table A.3). The other two predictors, library spending and published articles, did not cause similar behaviour. It was found that rate of software piracy and percent of internet access are highly correlated ($r = -.86, df = 26, p < 0.5$). Therefore, variance inflation factors (VIF) were calculated in order to assess if the model might be biased by colinearity, with an average $VIF > 1$ considered as an indicator for colinearity (Field, 2009, p. 224). The average $VIF = 2.65$ confirmed the presence of a bias attributable to colinearity (Table A.3 on page 61).

This could explain why the model that included all four variables had a smaller $R^2_{Adjusted}$, i.e., explained a smaller amount of the variance ($F(4, 22) = 6.597, p < .05, R^2 = .545, R^2_{Adjusted} = .463$). However, even the model with percent of internet access removed resulted with significant results ($F(3, 23) = 7.882, p < .05, R^2 = .507, R^2_{Adjusted} = .443$), and a significant value for the predictor published per staff ($\beta_1 = 7.871, t = 4.528, p < .05$) (Model 2 in Table A.3). This might indicate that piracy and internet access approximate the same unknown influence on downloads from SciHub.

3.6 Discussion

The analyses presented here set out to assess whether spending on libraries and or publishing output can explain the variation in SciHub usage across the EU. In addition, the degree of protection of intellectual property rights and the rate of software piracy were considered as possible factors that might influence SciHub use. Contrasting the volume of SciHub downloads against these variables was important in order to evaluate SciHub's role in alleviating issues of restricted access to literature in academia (Tenopir, Volentine, and King, 2012; Tenopir, Estelle, et al., 2015). More generally, however, although limited in scope, these analyses attempted to evaluate the potential of pirate libraries to ease human rights violations arising from the current publishing model, which, as discussed above, limits access to information (Byrne, 2007; Müller, 2010; Suber, 2012).

The main hypothesis tested herein proposed a relationship between investment in academic libraries and SciHub downloads by researchers, implying that SciHub should be used more often when conventional access to literature is more limited. The analysis, however, failed to recover such a relationship, resulting, instead, with essentially no interdependence between library expenditure per academic staff and the

frequency at which people download literature from SciHub (Figure 3.2 on page 38; Table A.1 on page 59). Although a negative relationship between SciHub downloads and library spending was an intuitive expectation, recovering such a pattern would have simply meant that resources such as SciHub matter mostly in countries that lag behind in investment on conventional libraries. However, the data showed that researchers in EU members with vastly different access-purchasing budgets can download articles from SciHub at similar rates (Figure 3.2, compare Finland, UK, and Poland). This suggested that 1) limitations to accessible literature persist despite substantial spending by some countries, and 2) the access provided by SciHub is of similar importance across countries regardless of the amounts invested in libraries.

The need for access to literature and the value of SciHub in alleviating the unanswered requirements is well exemplified by the substantial increase in the number of SciHub downloads across the EU over the six month period for which data were available (Figure 3.1 on page 31; Table 3.1 on page 32). Germany aside, even the smallest percent increase in SciHub downloads over this period was nearly 40% and amounted to over 13,000 articles (Table 3.1, Netherlands). Given the estimations that, once downloaded, articles can be re-shared up to 10 times (Tenopir, Estelle, et al., 2015), these volumes of downloads suggest that articles downloaded from pirate libraries have the potential to reach a substantial segment of academia. It is therefore plausible to propose that SciHub, and perhaps other similar services, are of considerable importance in alleviating restricted access to literature which has been detrimental towards fulfilment of the right to enjoy the benefits of scientific progress and its applications (REBSPA) and the right to information (Byrne, 2007; Müller, 2010).

This analysis rested on the simplifying assumptions that academics' needs for literature and journal subscription fees were the same across the EU. If UK's budget of about 7,000 USD per academic staff buys 70% of the needed articles (Online Computer Library Center, 2016), then Italy's 3,500 USD budget should buy 35%, resulting

in a greater access gap. It is of course possible that needs and cost vary across the EU. There is evidence that universities in the USA pay different amounts for subscriptions and the price can be influenced by several factors including existing relationships between a university and a publisher (Bohannon, 2014). Future research that takes such considerations into account should further refine these inferences regarding the relationship between investment in academic libraries and SciHub usage.

Publishing output of researchers across the EU, the second possible predictor of SciHub downloads tested here, was found to have a positive effect on SciHub usage (Figure 3.3 on page 39; Table A.2 on page 61). This result supported research activities as one of the drivers of SciHub downloads, and suggested that SciHub does not necessarily cater to ‘access needs’, in the sense that it is helpful only in places with little to no access, but to the ‘research needs’ of productive scholars who might experience access limitation to various degrees. Indeed, although published articles alone explained a relatively modest amount of the variance in SciHub downloads (Table A.2), the model indicated that up four articles were downloaded from SciHub for every published article.

That SciHub usage increases in countries with higher publishing outputs was expected to a degree. As researchers publish more, they need access to more literature, and it could be that they overcome shortcomings in access by downloading from SciHub. When considering publishing, however, it is also possible that the segment most severely affected by limited access are the highly productive researchers. Access and research needs, therefore, might interact to a certain degree to modulate the usage of SciHub across the EU. This interaction might be reflected in the fact that the most SciHub downloads per staff were recorded not in countries from Eastern Europe, where both access to literature and research capacity are generally lower, but in pre-2001 EU members including Greece, Portugal, Ireland, and Luxembourg (Figure 3.3). A possible explanation for this is that the austerity measures implemented in recent

years have had a detrimental effect on the available subscriptions in libraries. Thus, it might be that the socio-economic environment in these countries resulted in higher access limitations while the research capacity remained largely unchanged such that the newly expanded access gap was left to be complemented via alternative sources, perhaps including SciHub.

Although the roles of investment in libraries and publishing output were of primary interest in characterizing SciHub usage, aspects of the internet environment in a country, including the degree of enforcement of intellectual property law and overall internet access, are certainly expected to influence the frequency of SciHub downloads. The combined analysis of these variables suggested that downloads from SciHub were best explained when, in addition to investment and publishing, the model accounts for internet access (Table A.4 on page 63). However, counter-intuitively, internet access had a negative effect on SciHub downloads, suggesting that SciHub usage is greater in countries where internet access is available to less people. This was exemplified by Greece and Portugal, the two countries with most SciHub downloads, which simultaneously are among the countries with lowest internet access (Table A.6 on page 66). When considered independently from internet access, the commonness of software piracy – a proxy for the degree of enforcement of intellectual property law – had a positive effect on SciHub usage (Table A.4). This was an expected relationship, reflecting that pirate libraries were used more often if the opportunity was there and/or the potential legal cost of doing so was not a sufficient deterrent.

There were marginal differences between these two models, but when combined, internet access and software piracy cancelled each others explanatory power (Table A.4). A possible explanation for this is that these variables approximate a similar aspects or the socio-economic conditions in a country. Internet access might be an indicator of the development of the digital market in a country. Internet providers might expand their services if they know they can enable companies to sell more digital con-

tent (among other services). Such assurances, in turn, depend on the enforcement of intellectual property rights: companies can take advantage of the growing internet access if they are assured that free-riders, and thus negative effects on revenue, would be a minor concern. Thus, the improvement of internet access, relies on better enforcement of intellectual property rights which, in turn, reduces the levels of piracy, resulting with the observed cancelling effect of the two variables.

Notwithstanding, the best multivariate model managed to explain just below 50% of the variance of SciHub downloads by researchers. Thus, it is possible that percent internet access and software piracy were not the best choices for predictors of downloads from SciHub. Availability of SciHub across the EU that accounts for countries where the service is banned vs. those where it is available would have certainly been a better choice. However, to the best of the author's knowledge, such data are not available. The relatively poor explanatory power of the multivariate model could also result from the omission of other potentially important factors, including computer literacy among academics, age (older people might download less), or availability of library services off campus.

Overall, regardless of internet access and the degree of enforcement of intellectual property rights, including blocking SciHub.org, the general trend of SciHub downloads is upwards for all EU countries. A possible explanation here is that even countries with sound legal systems and comprehensive protection of intellectual property rights do not do enough to completely block SciHub. Whether this is because of technical reasons (not possible to block various domains or versions of the service), or because courts cannot react fast enough to the changing landscape of the internet, the end effect is poorer regulation of the access to SciHub, and as a by-product, improved access to otherwise inaccessible literature, both of which might have a positive effect on the fulfilment of REBSPA and the right to information, even at the slightest margins.

These findings have implications for the right to enjoy the benefits of scientific progress and its applications (REBSPA). Contributions to SciHub might be an example – the involvement of the scientific community in dissemination of scientific knowledge – one of the pleas of the Venice Statement (UNESCO, 2009) – and pirate libraries might have the potential to help EU members from the former Soviet bloc reach the standards of their western neighbours (Bodó, 2015). However, as argued above (Müller, 2010), the endmost goal of REBSPA is unrestricted distribution and application of scientific progress. If publishing output can be viewed as application of scientific progress, sharing literature by individuals might come short when it comes to fulfilling REBSPA. Since the obligation for development of scientific skills need for putting science into practice in order to fulfil REBSPA rests primarily on the states (UNESCO, 2009, p. 5), it appears that much more needs to be done to overcome the variety of differences that remain across EU states. It might be of little use to have access to the latest research if one does not have the means to apply it. Thus, mere availability of literature, whether through SciHub or elsewhere, might play a minor role in the efforts of Eastern European countries to achieve their western neighbours' levels of enjoyment of scientific progress and its applications – and by extension the broader realm of developing countries to the developed part of the world. The same can be said of the right to information in the context of scientific research: freedom of (scientific) expression might be valuable only to the extent that it is backed up by research skills.

In summary, spending at university libraries as a measure for access to scientific literature seems to be of minor importance in the usage of SciHub. Research activity of the academic community, measured through number of published articles, appears to be a better predictor, but is certainly not the only factor modulating SciHub use. Access to the internet and commonness of software piracy likely capture facets of the social surroundings with similarly strong, but opposing, influence over the usage of pirate libraries. These relationships, combined with volume of downloads and the continuing

trends of increase in SciHub usage, suggest that SciHub has a role in mitigating issues caused by limitations to literature access and, by extension, is of broader relevance for the struggle for human rights, including the right to information and REBSPA. However, the fulfilment of these rights cannot rely solely on improving access. Pirate libraries might complement poor access offered at university libraries, but this might not mean much if growth in research capacity and application of scientific discoveries is lagging.

Chapter 4: Conclusion

Following the stated objectives this work addressed the status and potential of pirate libraries through evaluation of whether access limitation violates human rights and whether the work of pirate libraries constitutes a struggle for advancement of human rights. In addition, as a measure of the severity of the problem, the extent of the literature access gap was assessed, i.e., how often do researchers run into road blocks related to obtaining literature. The dissertation then evaluated the usage of the currently most prominent pirate library, Sci-Hub.org, by researchers in the 28 member states of the European Union. This has not been investigated before, and speaks for the importance of SciHub in light of investment in purchasing access and the productivity of the scientific community.

Limitations to access to literature are widespread across EU states and even the most generous investors in journal subscriptions provide only two thirds of the literature needed for their scholars. Meanwhile, SciHub usage across the EU is on an upward trajectory. Some countries have reached up to 300% increase (Table 3.1 on page 32), and the EU as a whole posted as much as 382% increase in article downloaded from SciHub over a four month period (Figure 3.1 on page 31) Contrary to expectations, however, the amounts allocated to libraries were poor predictors of SciHub usage. Instead, the volume of downloads from SciHub was dependent on the research capacity as measured by publishing output, and was further influenced by availability of internet or commonness of software piracy. Thus, although access limitations certainly vary,

they nonetheless affect researchers throughout Europe, and scholars in countries with relatively high research capacity but poorer enforcement of intellectual property rights emerge as the most prolific users of SciHub. Taken together, these findings supported SciHub is an important resource, whose usage varies along political and economic lines, but is universally used by researchers to complement the conventional sources for obtaining literature. Given the ever-increasing pool of documents stored at pirate libraries' servers, SciHub and other services like it, might play a substantial role in ameliorating literature access limitations globally. More generally, and provided the current upward trends in usage are maintained in the longer run, pirate libraries have the potential to improve the fulfilment of the right to enjoy the benefits of scientific progress and its applications and the right to information.

The literature makes a case that lack of access to scientific literature is detrimental to human rights, and there is evidence that people involved in the work of pirate libraries share this idea. Although this assessment is shared by some researchers in the field (Cabanac, 2015; Gardner, C. C. and Gardner, G. J., 2017), the general view of academics might not necessarily mirror the ideas that restricting access violates human rights or that providing access is a human rights struggle (Tenopir, Volentine, and King, 2012; Volentine and Tenopir, 2013). This discrepancy might be a consequence of the research design of these studies, or it maybe the case is that the interviewed scholars did not consider access to literature a human rights issue. Thus, in contrast to academics who contribute to pirate libraries, and whose motivations echo the human rights issues discussed throughout this text, many other scholars might not view the various channels for access to literature as services that help solve a human rights problem. How academics perceive the interaction between access to literature and REBSPA or the right to (scientific) information might therefore be a fruitful ground for further research.

Beside access limitation stemming from the current publishing model and intellectual property rights regime, Pirate libraries can help overcome other limitations that conventional libraries face. Censorship, for example, is the most obvious limitation to access in countries where providing access is not just about money, but also about national security or protecting the ruling regime (Byrne, 2007). In this context, the 2.6 million SciHub downloads from Iran might offer an interesting research possibility. Questions such as do downloads in Iran reflect issues such as democracy, rule of law, women's rights, or religious freedoms can be a part of such a line of inquiry.

Overall, the implications of these results are only partly encouraging for the human rights struggles that pirate libraries represent. The universal access provided by the pirate libraries offers the opportunity for advancement of human rights, and foremost REBSPA. However, in the context of scientific research, the efforts to fulfil REBSPA and the right to information do not rely solely on providing access, but require improvements in the means to apply scientific progress both on and off research campuses. This is undoubtedly a complex endeavour that seems insurmountable without long-term commitment from governments and international organizations. Until then, the positive effects of pirate libraries are most likely be felt in already relatively healthy research communities or by small fractions of researchers in countries that lag behind in development of research and education. Thus, short of wholesale reform of the current publishing model and honest efforts in sharing not only the information, but the skills to acquire and apply scientific discoveries, comprehensive fulfilment of REBSPA and the right to information in the context of science appear to be out of reach.

Bibliography

- Armstrong, C., Beer, J., Kawooya, D., Prabhala, A. and Schonwetter, T., 2010. Summary and Conclusions. In: C. Armstrong, J. de Beer, D. Kawooya, A. Prabhala, and T. Schonwetter, eds. 2010. *Access to Knowledge in Africa: The Role of Copyright*. Claremont: UCT Press, p. 366. Available online at <https://www.idrc.ca/en/book/access-knowledge-africa-role-copyright>, (accessed on 20.06.2016).
- Bodó, B., 2015. Central and Eastern Europeans in Pirate Libraries. *Reference & User Services Quarterly*, 7.1, pp. 99–102. Available online at http://www.warsystems.hu/wp-content/uploads/sites/3/2015/06/7_Vise_Bodo-Balazs-2.pdf, (accessed on 16.06.2016).
- 2016a. In the Name of Humanity. *LIMN*, 6. Available online at <http://limn.it/in-the-name-of-humanity/>, (accessed on 16.06.2016).
- 2016b. In the Shadow of the Gigapedia. In: J. Karaganis, ed. 2016. *Shadow Libraries: Access to Educational Materials in the Digital Age*. New York: American Assembly. Electronic copy of the article obtained from the author.
- 2016c. The Russian Origins of the Online Shadow Library. In: J. Karaganis, ed. 2016. *Shadow Libraries: Access to Educational Materials in the Digital Age*. Columbia University: American Assembly. Electronic copy of paper obtained from the author.

- Bohannon, J., 2014. Secret bundles of profit. *Science*, 344.6190, pp. 1332–1333. Available online at <http://dx.doi.org/10.1126/science.344.6190.1332>, (accessed on 16.06.2016).
- 2016. Who’s downloading pirated papers? Everyone. *Science*, 352.6285, pp. 508–512. Available online at <http://dx.doi.org/10.1126/science.352.6285.508>, (accessed on 16.06.2016).
- Business Software Alliance, 2016. *Seizing Opportunity Through License Compliance: BSA Global Software Survey*. Available online at http://globalstudy.bsa.org/2016/downloads/studies/BSA_GSS_US.pdf (accessed 17.06.2016).
- Byrne, A., 2007. *The Politics of Promoting Freedom of Information and Expression in International Librarianship*. Plymouth: The Scarecrow Press, Inc., p. 226.
- Cabanac, G., 2015. Bibliogifts in LibGen? A study of a text-sharing platform driven by biblioleaks and crowdsourcing. *Journal of the Association for Information Science and Technology*, 67.4, pp. 874–884. Available online at <http://dx.doi.org/10.1002/asi.23445>, (accessed on 20.06.2016).
- Cenite, M., Wanzheng Wang, M., Chong, P. and Shimin Chan, G., 2009. More Than Just Free Content: Motivations of Peer-to-Peer File Sharers. *Journal of Communication Inquiry*, 33.3, pp. 206–221. Available online at <http://dx.doi.org/10.1177/0196859909333697>, (accessed on 20.06.2016).
- Dunn, G. A., Coiera, E. and Mandl, D. K., 2014. Is Biblioleaks Inevitable? *Journal of Medical Internet Research*, 16.4, e112. Available online at <http://dx.doi.org/10.2196/jmir.3331>, (accessed on 20.06.2016).
- Elbakyan, A. and Bohannon, J., 2016. *Data from: Who’s downloading pirated papers? Everyone*. *Dryad Digital Repository*. Available online at <http://dx.doi.org/10.5061/dryad.q447c> (accessed 16.06.2016).

- Elsevier, 2015. *Elsevier INC., Elsevier B.V., Elsevier LTD. v. Sch-Hub.org, Library Genesis Project, Alexandra Elbakyan, and John Does 1-99, Case 1:15-cv-04282-RWS, United States District Court, Southern District of New York*. Available online at <https://torrentfreak.com/images/elsevier-complaint.pdf> (accessed on 10.06.2016).
- Eurostat, 2016. *Classroom teachers and academic staff by education level, programme orientation, sex and age groups (educ_uoe_perp01)*. Available online at http://ec.europa.eu/eurostat/data/database?node_code=educ_uoe_perp01 (dataset last updated 30.05.2016, accessed 17.06.2016).
- Field, A. P., 2009. *Discovering statistics using SPSS*. 3rd ed. London: Sage Publications, p. 821.
- Foster, S. E., 2008. The Conflict between the Human Right to Education and Copyright. In: P. L. C. Torremans, ed. 2008. *Intellectual property and human rights: enhanced edition of Copyright and human rights*. Alphen aan den Rijn: Kluwer Law International, pp. 287–306.
- Gardner, C. C. and Gardner, G. J., 2017. Fast and Furious (at Publishers): The Motivations behind Crowdsourced Research Sharing. *College & Research Libraries*, Anticipated Publication Date: January 1, 2017. Available online at <http://crl.acrl.org/content/early/2016/02/25/crl16-840.abstract>, (accessed on 20.06.2016).
- Gordon, N., 2004. Human Rights as Being-Marginal-in-the-World. In: N. Gordon, ed. 2004. *From the Margins of Globalisation*. Oxford: Lexington Books, pp. 1–21.
- Helfer, L. R. and Austin, G. W., 2011. *Human Rights and Intellectual Property: Mapping the Global Interface*. Cambridge: Cambridge University Press, p. 550.
- Jaeger, P. T., Taylor, N. G. and Gorham, U., 2015. *Libraries, Human Rights, and Social Justice*. London: Rowman & Littlefield, p. 158.

- Jinha, A. E., 2010. Article 50 million: an estimate of the number of scholarly articles in existence. *Learned Publishing*, 23.3, pp. 258–263. Available online at <http://dx.doi.org/10.1087/20100308>, accessed on 18.06.2016).
- McNutt, M., 2016. Editorial: My love-hate of Sci-Hub. *Science*, 352.6285, p. 497. Available online at <http://dx.doi.org/10.1126/science.aaf9419>, (accessed on 16.06.2016).
- Müller, A., 2010. Remarks on the Venice Statement on the Right to Enjoy the Benefits of Scientific Progress and its Applications (Article 15 (1)(b) ICESCR). *Human Rights Law Review*, 10.4, pp. 765–784. Available online at <http://dx.doi.org/10.1093/hrlr/ngq033>, (accessed on 20.06.2016).
- Online Computer Library Center, 2016. *Global Library Statistics*. Available online at <https://www.oclc.org/global-library-statistics.en.html> (accessed 16.06.2016).
- Phenix, K. J. and Peña McCook, K., 2005. Human Rights and Librarians. *Reference & User Services Quarterly*, 45.1, pp. 23–26. Available online at <http://www.jstor.org/stable/20864437>, (accessed on 16.06.2016).
- R Development Core Team, 2016. *An Introduction to R*. Available online at <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf> (accessed 10.06.2016).
- Research4Life, 2016. *Access to Research in the Developing World*. Available online at <http://www.research4life.org/> (accessed 10.06.2016).
- Samek, T., 2007. *Librarianship and Human Rights: A Twenty-first Century Guide*. Oxford: Chandos Publishing, p. 200.
- Sample, I., 2012. *Harvard University says it can't afford journal publishers' prices*. Available online at <https://www.theguardian.com/science/2012/apr/24/harvard-university-journal-publishers-prices> (accessed on 10.07.2016).

- Scimago Journal & Country Rank, 2016. *Country Rankings*. Available online at <http://www.scimagojr.com/countryrank.php?year=2015> (accessed 16.06.2016).
- Suber, P., 2012. *Open Access*. London: MIT Press, p. 242.
- Swartz, A., 2008. *Guerilla Open Access Manifesto*. Available online at <http://ia600808.us.archive.org/17/items/GuerillaOpenAccessManifesto/Goamjuly2008.pdf> (accessed on 10.06.2016).
- Tenopir, C., Estelle, L., Haak, W., Allard, S., Christian, L., Nicholas, D., Watkinson, A., Woodward, H., Shepherd, P., Anderson, R. and Ali Saleh, S., 2015. The Secret Life of Articles: From Download Metrics to Downstream Impact. In: *Proceedings of the Charleston Library Conference*. Purdue University, pp. 1–11. Available online at <https://www.researchgate.net/publication/285491794> (accessed on 10.06.2016).
- Tenopir, C., Hughes, G., Christian, L., Allard, S., Nicholas, D., Watkinson, A., Woodward, H., Shepherd, P. and Anderson, R., 2014. To Boldly Go Beyond Downloads: How Are Journal Articles Shared and Used? In: *Proceedings of the Charleston Library Conference*. Purdue University, pp. 123–127. Available online at <http://dx.doi.org/10.5703/1288284315614> (accessed on 10.06.2016).
- Tenopir, C., Volentine, R. and King, D. W., 2012. UK Scholarly Reading and the Value of Library Resources: Summary Results of the Study Conducted Spring 2011. *insights*, 25.2, pp. 130–149. Available online at <http://insights.uksg.org/articles/10.1629/2048-7754.25.2.130/galley/158/download/>, (accessed on 20.06.2016).
- The World Bank, 2014. *Internet users (per 100 people)*. Available online at <http://data.worldbank.org/indicator/IT.NET.USER.P2> (accessed 17.06.2016).

- Torrent Freak, 2015a. *Elsevier Cracks Down on Pirated Scientific Articles*. Available online at <https://torrentfreak.com/elsevier-cracks-down-on-pirated-scientific-articles-150609/> (accessed on 10.06.2016).
- 2015b. *Sci-Hub, BookFi and LibGen Resurface After Being Shut Down*. Available online at <https://torrentfreak.com/sci-hub-and-libgen-resurface-after-being-shut-down-151121/> (accessed on 20.06.2016).
- 2015c. *Sci-Hub Tears Down Academia's "Illegal" Copyright Paywalls*. Available online at <https://torrentfreak.com/sci-hub-tears-down-academias-illegal-copyright-paywalls-150627/> (accessed on 10.06.2016).
- UNESCO, 2009. *Venice Statement on the Right to Enjoy the Benefits of Scientific Progress and its Applications*. Available online at http://www.aaas.org/sites/default/files/VeniceStatement_July2009.pdf (accessed on 10.06.2016).
- United Nations, 1948. *Universal Declaration of Human Rights*. Available online at <http://www.un.org/en/universal-declaration-human-rights/index.html> (accessed on 10.06.2016).
- 1966a. *International Covenant on Civil and Political Rights*. Available online at <http://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx> (accessed on 15.06.2016).
- 1966b. *International Covenant on Economic, Social and Cultural Rights*. Available online at <http://www.ohchr.org/EN/ProfessionalInterest/Pages/CESCR.aspx> (accessed on 10.06.2016).
- 1993. *Vienna Declaration and Programme of Action*. Available online at <http://www.ohchr.org/EN/ProfessionalInterest/Pages/Vienna.aspx> (accessed on 10.07.2016).

Volentine, R. and Tenopir, C., 2013. Value of academic reading and value of the library in academics' own words. *Aslib Proceedings*, 65.4, pp. 425–440. Available online at <http://dx.doi.org/10.1108/AP-03-2012-0025>, (accessed on 20.06.2016).

Ware, M. and Mabe, M., 2015. *The STM Report: An overview of scientific and scholarly journal publishing*. International Association of Scientific, Technical and Medical Publishers. Available online at http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf (accessed on 20.07.2016).

Appendix A: SciHub Data Analysis

Table A.1: Linear regression between number of SciHub downloads per academic staff and library spending per staff across the EU. All cases (1) and without outliers (2).

	<i>Dependent variable:</i>	
	SciHub downloads per staff	
	(1)	(2)
Library spending per staff	−0.0004 (0.0003) t = −1.222 p = .233	−0.0004 (0.0003) t = −1.442 p = .163
Intercept	5.305*** (1.038) t = 5.109 p = .00003	4.700*** (0.815) t = 5.765 p = .00001
Observations	27	26
R ²	.056	.080
Adjusted R ²	.019	.041
Residual Std. Error	4.378 (df = 25)	3.384 (df = 24)
F Statistic	1.494 (df = 1; 25) (p = .233)	2.079 (df = 1; 24) (p = .163)
<i>Note:</i>	*p<.1; **p<.05; ***p<.01	

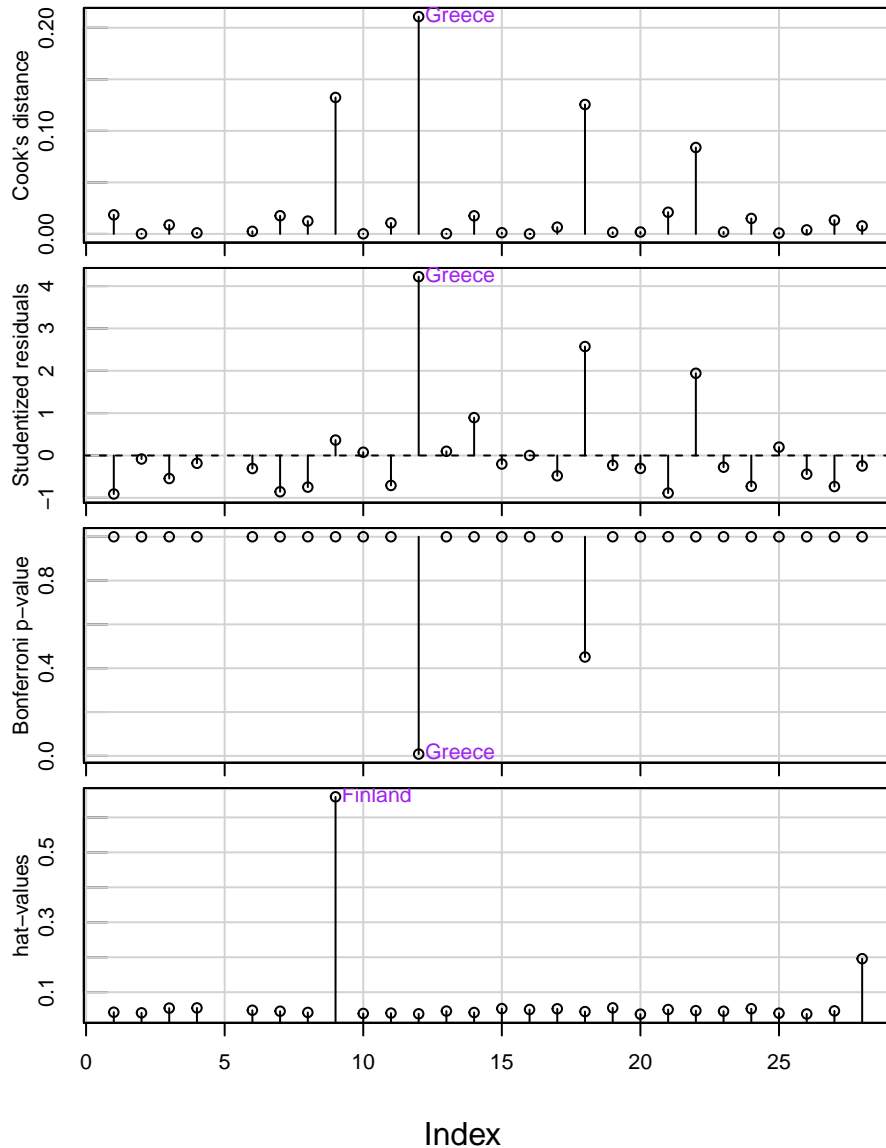


Figure A.1: Diagnostic plot for the regression of SciHub downloads per staff by library spending per staff. Shown are the Cook's distances (panel one), standardized (Studentized) residuals (panel two), Bonferroni corrected Studentized residuals (panel three), and the leverage data points (panel four). After Bonferroni correction for multiple comparisons only Greece has an (adjusted) p value $< .05$. Greece is simultaneously an influential point (Cook's distance). By being both influential and an outlier, Greece might have an effect on the regression model.

Table A.2: Linear regression between number of SciHub downloads per academic staff and number of published papers per staff across the EU. All cases (1) and without outliers (2).

	<i>Dependent variable:</i>	
	SciHub downloads per staff	
	(1)	(2)
Pub. papers per staff	4.346** (1.785) t = 2.435 p = .023	3.399** (1.414) t = 2.404 p = .024
Intercept	1.260 (1.518) t = 0.830 p = .415	1.472 (1.188) t = 1.239 p = .227
Observations	28	27
R ²	.186	.188
Adjusted R ²	.154	.155
Residual Std. Error	4.019 (df = 26)	3.142 (df = 25)
F Statistic	5.928** (df = 1; 26) (p = .023)	5.781** (df = 1; 25) (p = .024)

Note: *p<.1; **p<.05; ***p<.01

Table A.3: Variance inflation factors for model with Downloads per staff as dependent variable.

	Value
Published per staff	1.44
Library spending per staff	1.29
Rate of software piracy	3.86
Percent of internet access	3.99
Average VIF	2.65

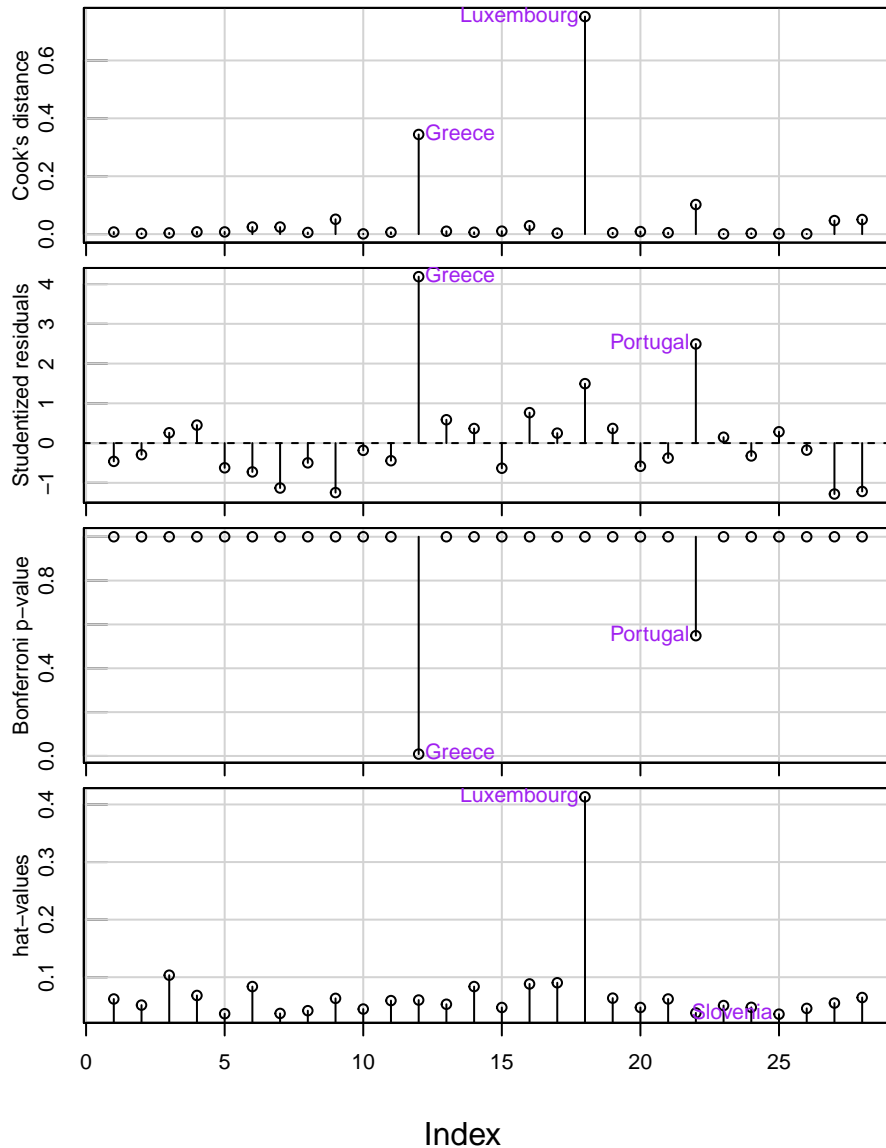


Figure A.2: Diagnostic plot for the regression of SciHub downloads per staff by number of published papers per staff. Shown are the Cook's distances (panel one), standardized (Studentized) residuals (panel two), Bonferroni corrected Studentized residuals (panel three), and the leverage data points (panel four). After Bonferroni correction for multiple comparisons only Greece has an (adjusted) p value $< .05$. In addition to Luxembourg, Greece is also an influential point (Cook's distance). By being both influential and an outlier, Greece might have an effect on the regression model.

Table A.4: Multivariate regression of SciHub downloads per academic staff by the number of published papers per staff, library spending per staff, rate of software piracy, and percent of internet access across the EU (1); rate of software piracy removed from the model (2); percent of internet access removed from the model (3).

	<i>Dependent variable:</i>		
	SciHub downloads per staff		
	(1)	(2)	(3)
Pub. papers per staff	8.212*** (1.725) t = 4.761 p = .0001	7.871*** (1.738) t = 4.528 p = .0002	8.082*** (1.680) t = 4.809 p = .0001
Library spending per staff	-0.001* (0.0003) t = -1.849 p = .078	-0.001* (0.0003) t = -2.066 p = .051	-0.001* (0.0003) t = -1.903 p = .070
Rate of software piracy	4.886 (9.205) t = 0.531 p = .601	14.762** (5.784) t = 2.552 p = .018	
Percent of internet access			-18.890*** (6.440) t = -2.933 p = .008
Intercept	8.881 (11.090) t = 0.801 p = .432	-5.660* (3.095) t = -1.829 p = .081	14.246*** (4.493) t = 3.171 p = .005
Observations	27	27	27
R ²	.545	.507	.540
Adjusted R ²	.463	.443	.479
Residual Std. Error	3.239 (df = 22)	3.299 (df = 23)	3.188 (df = 23)
F Statistic	6.597*** (df = 4; 22) (p = .002)	7.882*** (df = 3; 23) (p = .001)	8.983*** (df = 3; 23) (p = .0005)

Note:

*p<.1; **p<.05; ***p<.01

Table A.5: Linear regression between number of SciHub downloads per student and annual expenditure per tertiary level student across the EU. All cases (1) and without outliers (2).

	Dependent variable:	
	SciHub downloads per student	
	(1)	(2)
Expenditure per student	-0.00001* (0.00000) t = -1.750 p = .096	-0.00001* (0.00000) t = -1.897 p = .074
Intercept	0.299*** (0.050) t = 6.004 p = .00001	0.280*** (0.043) t = 6.485 p = .00001
Observations	22	21
R ²	.133	.159
Adjusted R ²	.089	.115
Residual Std. Error	0.122 (df = 20)	0.105 (df = 19)
F Statistic	3.061* (df = 1; 20) (p = .096)	3.600* (df = 1; 19) (p = .074)

Note:

*p<.1; **p<.05; ***p<.01

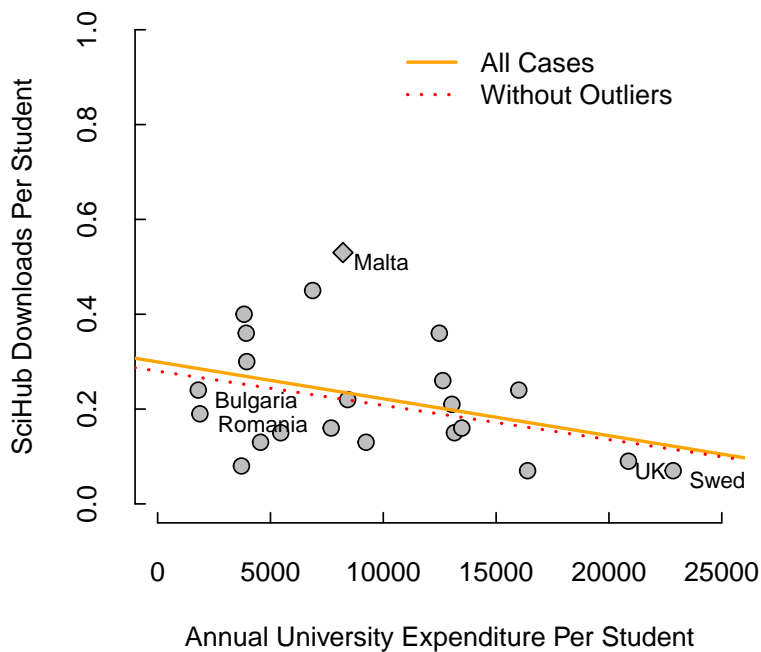


Figure A.3: Linear regression of SciHub downloads per student and annual expenditure per tertiary-level student across the EU. All cases (full line) and without (dashed line) the outlier country Malta.

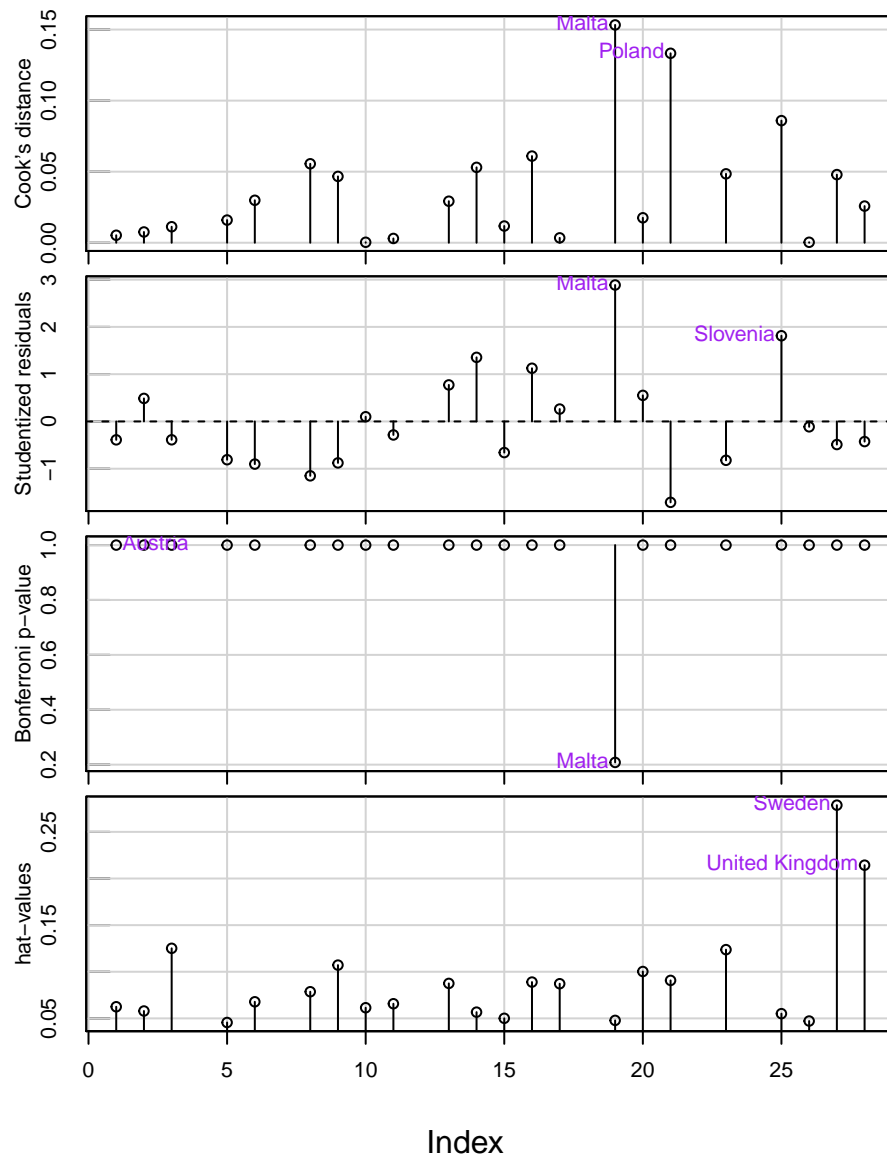


Figure A.4: Diagnostic plot for the regression of SciHub downloads per student by annual expenditure per tertiary-level student. Shown are the Cook's distances (panel one), standardized (Studentized) residuals (panel two), Bonferroni corrected Studentized residuals (panel three), and the leverage data points (panel four). After Bonferroni correction for multiple comparisons only Malta has an (adjusted) p value $< .05$. In addition to Poland, Malta is also an influential point (Cook's distance). By being both influential and an outlier, Malta might have an effect on the regression model.

Table A.6: Dataset Used in Analysis

EU Code ^a	Country	Downloads Per Staff ^b	Library Spending Per Staff ^c	Published Papers Per Staff ^d	Percent Of Internet Access ^e	Rate Of Software Piracy ^f	Downloads Per Student ^g	Annual Expenditure Per Student ^h
15	Austria	1.05	855.72	0.37	0.81	0.21	0.15	13153
15	Belgium	4.53	1016.80	1.02	0.85	0.23	0.26	12636
13	Bulgaria	2.92	90.87	0.15	0.55	0.60	0.24	1803
13	Croatia	4.48	64.05	0.33	0.69	0.51	0.45	NA
13	Cyprus	1.69	NA	0.67	0.69	0.45	0.13	9236
13	Czech Republic	3.79	432.55	1.23	0.80	0.33	0.15	5455
15	Denmark	0.39	3067.50	0.82	0.96	0.22	0.04	NA
13	Estonia	1.71	904.56	0.56	0.84	0.42	0.13	4560
15	Finland	1.29	12338.43	1.11	0.92	0.24	0.07	16395
15	France	4.66	2436.76	0.95	0.84	0.34	0.21	13041
15	Germany	1.20	2580.18	0.39	0.86	0.22	0.16	13479
15	Greece	18.75	1417.09	1.09	0.63	0.63	0.42	NA
13	Hungary	5.50	587.61	0.44	0.76	0.38	0.36	3926
15	Ireland	8.04	2760.98	1.23	0.80	0.32	0.36	12483
15	Italy	3.01	3523.74	0.98	0.62	0.45	0.16	7690
13	Latvia	5.18	306.11	0.22	0.76	0.49	0.40	3831
13	Lithuania	3.15	194.15	0.21	0.72	0.51	0.30	3949
15	Luxembourg	14.97	702.27	2.12	0.95	0.19	1.76	NA
13	Malta	4.28	55.38	0.36	0.73	0.44	0.53	8213
15	Netherlands	3.20	1909.32	0.98	0.93	0.24	0.24	16004
13	Poland	1.38	308.27	0.37	0.67	0.48	0.08	3714
15	Portugal	12.97	508.98	0.63	0.65	0.29	1.20	NA
13	Romania	3.85	621.64	0.46	0.54	0.60	0.19	1873
13	Slovakia	2.10	181.70	0.49	0.80	0.36	0.14	NA
13	Slovenia	5.71	1138.19	0.76	0.72	0.43	0.45	6869
15	Spain	2.77	1514.20	0.51	0.76	0.44	0.22	8421
15	Sweden	0.88	3134.32	1.05	0.93	0.21	0.07	22844
15	United Kingdom	1.44	7145.53	1.12	0.92	0.22	0.09	20864

^aIndicates membership in the EU, before or after eastern expansion of 2001

^bCalculated as number of SciHub downloads in a country divided by the number of academic staff

^cCalculated as a country's spending on academic libraries divided by number of academic staff

^dCalculated as total number of papers published by a country divided by the number of academic staff

^eInternet users (per 100 people)

^fNumber of Unlicensed Software Units divided by Total Software Units Installed

^gCalculated as the number of SciHub downloads in a country divided by the number of students

^hCalculated as the total annual expenditure on tertiary education in a country divided by the number of students