

Legal Research in Search of Attention: A Quantitative Assessment

Mathias M Siems*

*Final version published in
King's Law Journal, Volume 27, 2016, pp. 170-187*

Abstract. In today's world it is easy to make research publicly available by putting it online. But this improved availability raises the question how to produce research that actually gets attention. Bibliometrics can contribute to this debate. Based on a sample of 1107 papers of SSRN's Legal Scholarship Network, this article finds that a short title, a top-20 university affiliation, US authorship, and writing about topics of corporate law and international law have a positive effect on downloads and/or abstract views. The article also reflects on the implications of these findings, in particular how they may be related to contentious attempts to identify what is 'good' legal research through metrics and peer review.

Keywords: legal research; higher education; impact metrics; bibliometrics; REF

JEL Codes: C80, I23, K00

* Professor of Commercial Law, Durham University and Research Associate, Centre for Business Research, University of Cambridge. I thank Rob van Gestel, an anonymous referee and the participants of the Durham Law School Research Seminar Series for helpful comments. The usual disclaimer applies.

INTRODUCTION

Academic staff are expected to produce research and, in today's world, it is then also straightforward to make this research publicly available by putting it online. Yet, this improved availability raises the question how to produce research that actually gets attention. The high degree of non-citation of research is frequently discussed in the higher education literature.¹ Legal research is no exception: for example, it has been found that 43% of articles in Lexis-Nexis are never cited.²

Bibliometrics can contribute to this debate. Specifically, the Social Science Research Network (SSRN) is a good platform to test which research is more or less appealing. In the study reported in this article, 1107 papers of SSRN's Legal Scholarship Network were analysed in order to identify the main determinants of SSRN downloads, abstract views, and downloads per abstract views. This analysis fills a gap in the growing literature that deals with the impact of published research.³ It is also suggested that examining SSRN is important because its open nature reflects the general trend from offline publications in domestic journals to global availability of publications online.⁴

The substantive contribution of the study is that it can show whether particular types of legal scholarship are regarded as more attractive, how much formalities matter, and how much scholars from elite universities have an advantage. Understanding those factors is important because it may demonstrate whether or not SSRN indicators may be able to capture the quality of legal publications. The present research therefore also contributes to the discussion about assessment systems of research, such as the choice between peer-review systems, journal rankings, and citation metrics.

The structure of this article is as follows. The next section explains the theoretical framework underlying this article in more detail, followed by a section that presents the data collection and outlines the dependent and explanatory variables used in this study. The subsequent section reports and discusses the regression results. The final section concludes.

¹ See references in Liming Liang, Zhen Zhong, and Ronald Rousseau, 'Uncited Papers, Uncited Authors and Uncited Topics: A Case Study in Library and Information Science' (2015) 9 *Journal of Informetrics* 50.

² Thomas A Smith, 'The Web of Law' (2007) 44 *San Diego Law Review* 309, 335-6.

³ See the references in following section.

⁴ *Ibid* for references.

The reception of scholarship: bibliometrics and research assessments

The reception of works of scholarship can be conceptualised as being dependent on both internal and external factors. Internal factors are those over which the author is in full control. Some of them may be formal ones: for example, it may matter whether a piece is of sufficient length to make readers interested in it. In substance, ideally, one would expect that good quality research is rewarded.⁵ As regards the role of external factors, for example, it may be the case that writing about popular academic topics and adopting mainstream methods may lead to a large audience.⁶ However, it can also be suggested that it is helpful to adopt an interdisciplinary, international or practical perspective in order to reach beyond the audience of domestic scholarship.

Bibliometric indicators by major publishers, such as Scopus and the Thomson Reuters Web of Science (formerly ISI Web of Knowledge),⁷ use citation data in order to capture the reception of published research. The importance of those indicators cannot be underestimated. Citation-based journal impact factors are widely used in many disciplines⁸ (though less frequent in law⁹). The citations of individual pieces of research can also be important for academics whose universities rely on metrics to measure research impact (eg, through the ‘h-index’).¹⁰ In addition, aggregates of citation scores are often a component of university rankings, with the QS ranking also using (incomplete) citation data for the sub-ranking of top universities in law.¹¹

⁵ For further discussion about ‘good quality’ in legal research see text to n 73, below.

⁶ On the pressure to follow the ‘dominant paradigm’ see, eg, Pierre Schlag, ‘Spam Jurisprudence, Air Law, and the Rank Anxiety of Nothing Happening (A Report on the State of the Art)’ (2009) 97 *Georgetown Law Journal* 803.

⁷ See www.scopus.com and www.webofscience.com.

⁸ See Derek R Smith, ‘Impact Factors, Scientometrics and the History of Citation-Based Research’ (2012) 92 *Scientometrics* 419.

⁹ However, <http://lawlib.wlu.edu/LJ/> includes such information for a wide range of law journals.

¹⁰ See, eg, www.harzing.com/pop_hindex.htm.

¹¹ See www.topuniversities.com/university-rankings/university-subject-rankings/2015/law-legal-studies.

The academic literature of many fields has examined the internal and external factors that can explain those citations.¹² There has also been some research on the relationship between journal article downloads and citations.¹³ With respect to legal journals, research on possible determinants of citations has mainly scrutinised US law reviews,¹⁴ which have the advantage that they are fully available electronically (through Westlaw) and therefore can be searched easily.

The reception of scholarship and its determinants are also themes of government-sponsored research assessments of universities. A prominent model is that of the UK Research Assessment Exercises (RAEs), now rebranded as Research Excellence Framework (REF), which has also been adopted, with some modifications, in other countries such as Australia, New Zealand, Hong Kong, Italy and the Netherlands.¹⁵

The most recent exercise, the REF 2014, included an explicit peer-review based assessment of cases studies that demonstrated impact *beyond* academia.¹⁶ The main element, however, is still the quality of research. Here, it was controversial whether the impact *within* academia can be seen as a meaningful proxy. Some have suggested that bibliometric indicators are cheaper and more precise than a peer review process,¹⁷ while a quantitative assessment of the previous RAE found that citation-based indicators are

¹² References in John Mingers and Fang Xu, 'The Drivers of Citations in Management Science Journals' (2010) 205 *European Journal of Operational Research* 422.

¹³ Vincente P Guerrero-Bote and Félix Moya-Anegón, 'Relationship Between Downloads and Citations at Journal and Paper Levels, and the Influence of Language' (2014) 101 *Scientometrics* 1043; Carsten Nieder, Astrid Dalhaug and Gro Aandahl, 'Correlation between Article Download and Citation Figures for Highly Accessed Articles from Five Open Access Oncology Journals', *SpringerPlus* 2013, 2:261, available at www.springerplus.com/content/2/1/261.

¹⁴ Eg, Ian Ayres and Fredrick E Vars, 'Determinants of Citations to Articles in Elite Law Reviews' (2000) 29 *The Journal of Legal Studies* 427; Alfred L Brophy, 'Law [Review]'s Empire: The Assessment of Law Reviews and Trends in Legal Scholarship' (2006) 39 *Connecticut Law Review* 101; Alfred L Brophy, 'The Signaling Value of Law Reviews: An Exploration of Citations and Prestige' (2009) 36 *Florida State University Law Review* 229.

¹⁵ See generally, eg, Giovanni Abramo and Ciriaco Andrea D'Angelo, 'Evaluating Research: From Informed Peer Review to Bibliometrics' (2011) 87 *Scientometrics* 499; Richard Whitley and Jochen Gläser (eds), *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems* (Dordrecht, Springer, 2007); Aldo Geuna and Ben R Martin, 'University Research Evaluation and Funding: An International Comparison' (2003) 41 *Minerva* 277.

¹⁶ See the section on 'impact' in *Assessment Framework and Guidance on Submissions* (HEFCE 2011), available at www.ref.ac.uk/pubs/2011-02/.

¹⁷ Eg, Jim Taylor, 'The Assessment of Research Quality in UK Universities: Peer Review or Metrics?' (2011) 22 *British Journal of Management* 202; Patrick Dunleavy, 'The Research Excellence Framework is lumbering and expensive' (9 June 2011) available at <http://blogs.lse.ac.uk/politicsandpolicy/archives/10909>.

not reliable in the social sciences and humanities.¹⁸ A pilot exercise of the REF 2014 considered the use of bibliometric indicators, but then took the view that ‘citation information is not sufficiently robust to be used formulaically or as a primary indicator of quality’.¹⁹ Citation data were therefore only included as an additional element for some scientific disciplines, not law.²⁰ This cautious approach is continued in a review of the role of metrics in research assessment in preparation for the next REF.²¹

The research assessments in Australia also use citation data for some disciplines only.²² The Australian Research Council previously also suggested another approach, namely, to use peer-review based journal rankings in order to evaluate the quality of research outputs. However, after considerable criticism, these rankings were abandoned.²³ Journal rankings have also been frequent but controversial topics of academic research. In business and economics, where impact-based journal rankings play a major role, it has recently been suggested that those rankings are not reliable proxies for quality.²⁴ With respect to law, the results of the UK’s RAE/REFs have been used to rank law journals,²⁵ while many legal scholars tend to be sceptical of law journal rankings.²⁶

Overall, it can be seen that journal-based bibliometrics and research assessment exercises have, to some extent, been interested in the attention that legal research receives.

¹⁸ O Mryglod, R Kenna, Y Holovatch and B Berche, ‘Comparison of a Citation-Based Indicator and Peer Review for Absolute and Specific Measures of Research-group Excellence’ (2013) 97 *Scientometrics* 767. See also Rob van Gestel and Jan Vranken, ‘Assessing Legal Research: Sense and Nonsense of Peer Review versus Bibliometrics and the Need for a European Approach’ (2011) 12 *German Law Journal* 901.

¹⁹ See www.ref.ac.uk/about/background/bibliometrics/.

²⁰ See www.ref.ac.uk/about/guidance/citationdata/.

²¹ See www.hefce.ac.uk/rsrch/metrics/.

²² See www.arc.gov.au/era/era_2015/era_2015.htm.

²³ See <https://research.unsw.edu.au/excellence-research-australia-era-outlet-ranking>.

²⁴ Ismael Rafols, Loet Leydesdorff, Alice O’Hare, Paul Nightingale, and Andy Stirling, ‘How Journal Rankings Can Suppress Interdisciplinary Research: A Comparison Between Innovation Studies and Business & Management’ (2012) 41 *Research Policy* 1262; Frederic S Lee, Xuan Pham, and Gyun Gu, ‘The UK Research Assessment Exercise and the narrowing of UK economics’ (2013) 37 *Cambridge Journal of Economics* 693 (referring both to journal rankings and the RAEs).

²⁵ For the RAE 2001: Kevin Campbell, Alan Goodacre and Gavin Little, ‘Ranking of United Kingdom Law Journals: An Analysis of the Research Assessment Exercise 2001 Submissions and Results’ (2006) 33 *Journal of Law and Society* 335. For the REF 2014: Mathias Siems, ‘Law Journal Ranking (based on REF 2014)’ (19 January 2015), available at <http://siemslegal.blogspot.co.uk/2015/01/law-journal-ranking-based-on-ref-2014.html>.

²⁶ See the discussion in Rob van Gestel, ‘Sense and Non-sense of a European Ranking of Law Schools and Law Journals’ (2015) 35 *Legal Studies* 165.

But the following will explain how researching SSRN has some advantages over those former approaches.

Researching SSRN: benefits and idiosyncrasies

The Social Science Research Network (SSRN) was founded in 1994 and is, in its own words, ‘devoted to the rapid worldwide dissemination of social science research’.²⁷ Everyone can upload papers on SSRN and almost all of the papers are free to download. SSRN is not the only way legal scholars make their research freely available. For example, Academia.edu and Researchgate.net now provide similar services, and some universities make their faculty scholarship available in online repositories; however, this still leaves SSRN with a ‘first-mover advantage’ and, at least at the moment, with a dominant position in the market for freely available academic papers.

Understanding the determinants of SSRN downloads can be of interest since the SSRN download ranks of universities and academics have at least some practical relevance. Highly ranked universities often advertise their good performance,²⁸ presumably since this is seen as helpful for marketing purposes. In US universities, but also elsewhere, it is not uncommon that researchers include their SSRN author ranks in their CVs, presumably since this can play a role for job applications and promotions.

More fundamentally researching SSRN downloads is valuable since SSRN can be seen as paradigmatic for the evolution of academic publications. The traditional model was that of a small number of articles published in a small number of journals and only accessible to a small number of persons (namely, academics working in this field who had access to the hardcopy of the journal). Over time, more journals have emerged and now the vast majority of journals are available online. Thus, today, published journal articles have a much more diverse readership: academic readers may find articles from any field using ‘Google scholar’ and can then access them via their institutional journal subscriptions. Members of the public also come across academic journal articles more often (with ‘Google’, ‘Bing’ etc.) and, if this article or journal has been made ‘open ac-

²⁷ SSRN, FAQs (2014), available at www.ssrn.com/update/general/ssrn_faq.html.

²⁸ See, eg, [/www.tilburguniversity.edu/about/schools/law/](http://www.tilburguniversity.edu/about/schools/law/) and <http://today.law.harvard.edu/thirteen-harvard-law-school-faculty-on-ssrns-100-most-cited-law-school-professors/>.

cess’,²⁹ they are then also able to download this piece of research. As a result, journal publications have gradually become similar to the open model of SSRN.

The open and democratic nature of SSRN provides further benefits. Thinking about legal scholarship as a ‘market for good ideas’, and acknowledging that the ‘wisdom of the crowd’ has some truth to it,³⁰ SSRN indicators may reward the supply of original research.³¹ The general availability of SSRN papers is particularly helpful for research on legal publications. The Legal Scholarship Network (LSN) is one of the specialised subject area networks of SSRN. Analysing papers of this network is valuable because journal-based services such as the Web of Science or Scopus³² exclude many law journals.³³ It is also possible to consider the advantages of a large audience of readers. While metrics based on peer review and journal rankings tend to focus on the domestic mainstream,³⁴ SSRN indicators should be better able to capture interdisciplinary and international legal research. It also considers that some papers may be particularly helpful for purposes of teaching. In addition, since SSRN papers are freely available, papers should benefit from any impact they have beyond academia.

The SSRN study of the present article also fills a gap in the literature. So far, research on SSRN papers has, for example, analysed its citation network,³⁵ the occurrence of self-downloads,³⁶ and whether SSRN downloads should be used to rank law

²⁹ For the discussion see, eg, the various articles available at www.theguardian.com/science/open-access-scientific-publishing.

³⁰ Popularised by James Surowiecki, *The Wisdom of Crowds* (New York: Doubleday, 2004).

³¹ To be sure, this is eventually an empirical question, see text to n 73, below.

³² See n 7.

³³ See, eg, Thed Van Leeuwen, ‘Bibliometric Research Evaluations, Web of Science and the Social Sciences and Humanities: A Problematic Relationship?’ (2013) 2 *Bibliometrie – Praxis und Forschung* 8, available at www.bibliometrie-pf.de/article/viewFile/173/218 (less than 40% coverage of law journals in Web of Science).

³⁴ See references n 24 (for business studies and economics). See also ‘Research that doesn’t belong to single subject area is deemed “too risky”’ (Guardian, 21 Nov 2013), available at www.theguardian.com/higher-education-network/blog/2013/nov/21/interdisciplinary-research-ref-submission-university.

³⁵ Jeven D West, Michael C Jensen, Ralph J Dandrea, Gregory J Gordon and Carl T Bergstrom, ‘Author-level Eigenfactor Metrics: Evaluating the Influence of Authors, Institutions, and Countries within the Social Science Research Network Community’ (2013) 64 *Journal of the American Society for Information Science and Technology* 787.

³⁶ Benjamin G Edelman and Ian Larkin, ‘Social Comparisons and Deception Across Workplace Hierarchies: Field and Experimental Evidence’, Harvard Business School NOM Unit Working Paper No 09-096 (2013), available at <http://ssrn.com/abstract=1346397>.

schools.³⁷ Only one paper seems to have examined one of the factors that may determine SSRN downloads, namely, the influence of blog posts.³⁸

Finally, it needs to be noted that SSRN has some idiosyncrasies. There is a well known example that shows that SSRN downloads may reward provocative themes and titles.³⁹ Despite its global availability, there is also bound to be certain preference for legal scholarship published in English, not least since the SSRN website is only available in English (though it allows uploading papers in other languages). Since SSRN has its origins in US academia, there is also likely to be a bias for research on US law, going beyond the size of the US market for legal scholarship. This factor may then also have an impact on the preferred methods given that in the US, but not for instance in Europe, the interdisciplinary approach to legal scholarship is said to have won the day.⁴⁰ The following will therefore account for and discuss these specific features of SSRN.

VARIABLES AND DESCRIPTIVE STATISTICS

Dependent variables

Some citation studies aim to examine the lifetime of citations, in particular whether and when article citations peak.⁴¹ By contrast, the present article aims to keep the time of the

³⁷ Bernard S Black and Paul L Caron, 'Ranking Law Schools: Using SSRN to Measure Scholarly Performance' (2006) 81 *Indiana Law Journal* 83; Lawrence A Cunningham, 'Commentary, Scholarly Profit Margins and the Legal Scholarship Network: Reflections on the Web' (2006) 81 *Indiana Law Journal* 271; Theodore Eisenberg, 'Commentary, Assessing the SSRN-Based Law School Rankings' (2006) 81 *Indiana Law Journal* 285.

³⁸ Paul Ohm, 'Do Blogs Influence SSRN Downloads? Empirically Testing the Volokh and Slashdot Effects', University of Colorado Law Legal Studies Research Paper No 07-15 (2007), available at <http://ssrn.com/abstract=980484>.

³⁹ Cf Christopher M Fairman, 'Fuck' (2007) 28 *Cardozo Law Review* 1711 (SSRN version at <http://ssrn.com/paper=896790>); Christopher M Fairman, 'Fuck and Law Faculty Rankings', Ohio State Public Law Working Paper 2007 No 91, available at <http://ssrn.com/abstract=971103>.

⁴⁰ See Mathias Siems and Daithi Mac Sithigh, 'Mapping Legal Research' (2012) 71 *Cambridge Law Journal* 651. But, in the US, this trend has also been criticised: eg, Brent E Newton, 'Preaching What They Don't Practice: Why Law Faculties' Preoccupation with Impractical Scholarship and Devaluation of Practical Competencies Obstruct Reform in the Legal Academy' (2010) 62 *South Carolina Law Review* 105; Harry T Edwards, 'The Growing Disjunction between Legal Education and the Legal Profession' (1992) 91 *Michigan Law Review* 34.

⁴¹ Eg, Ayres and Vars (n 14).

respective data relatively constant. This has the advantage that it avoids the problem of a non-stationary time series, namely that SSRN may have become less or more popular in recent years. The idea was therefore to collect data of freely available LSN-SSRN papers⁴² that were uploaded at approximately the same time, and were still available two years later.

Initially, it was analysed how many downloads and abstract views the LSN papers posted between 1 and 12 January 2012 had on 12 January 2014. But subsequently it was considered that it may be the case that the January papers could be somehow unusual due to the post-holiday period. Thus, in a second step, it was examined how many downloads and abstract views the LSN papers posted between 13 and 25 October 2012 had on 25 October 2014. This led to 553 papers for the January and 554 for the October data, thus in total 1107 papers. This was seen as a good sample size as most of the explanatory variables (explained below) required hand-collected data. The subsequent regressions will also control for slight differences in the number of days each of the papers has been online (735 to 747) as well as potential differences between the January and October data.

⁴² Papers were excluded which are not freely available or where only the abstract is provided given that the following also aims to examine the relationship between abstract views and downloads.

Figure 1: Box plot of downloads for January (left) and October (right)

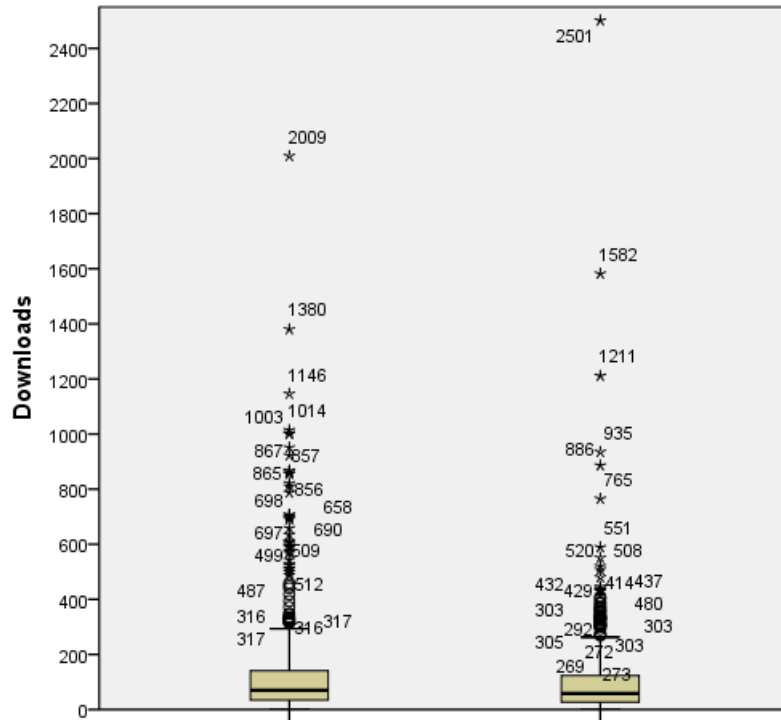


Figure 1 displays the January and October data for SSRN downloads in a box plot.⁴³ It can be seen that the shape is similar: a number of papers with more than 300 downloads but the majority of papers with considerably lower numbers. The actual data show that the January 2012 papers have slightly more downloads than the October ones: the median is 70 for January and 58 for October, and the means are 136.58 and 108.54, while the interquartile range (the upper and lower limits of the box) is more similar: 108 and 97.⁴⁴

Table 1: Descriptive Statistics of Dependent Variables (n: 1107)

	Mean	Median	Std. Deviation	Minimum	Maximum
Abstract views	564.62	369	627.999	17	7321
Paper downloads	122.55	63	188.844	2	2501
Downloads per view	0.1874	0.1711	0.12	0.01	1.23

⁴³ In a box plot the top and bottom of the ‘box’ indicate the third and first quartile (the interquartile range, ie the middle half of the data points); the horizontal line within this box indicates the median; the ‘whiskers’ indicate the maximum and minimum.

⁴⁴ It is beyond the scope of this article to explore why this may be the case given that the subsequent regressions (see Table 3, below) aim to control for this factor.

Table 1 reports the abstract views and paper downloads for all papers of this study. These two numbers are strongly correlated (0.852), but the following will also examine whether there may be reasons, say, related to the abstract of the paper, that may matter for the download but not the abstract view (or vice versa). In addition, a third dependent variable was used: downloads per abstract views. This was done in order to examine more precisely what accounts for a download of a paper *given* that someone already had a look at its abstract page. Such information may be particularly interesting if it is assumed that the mere number of visits to the SSRN abstract pages may be influenced by factors that are not entirely rational, say, the preference of Google or other search engines for certain words and phrases.

By contrast, this article is not based on citations of the papers in question. Since 2008 SSRN provides a service called CiteReader.⁴⁵ However, the problem is that it only tracks citations in the SSRN database; thus, at the moment, it is only of limited value.⁴⁶ It may also be suggested to identify the citations of SSRN papers with Google Scholar but, here too, the coverage would be incomplete since many books and traditional law journals are not covered. Moreover, this article is based on the motivation that information on abstract views and downloads is interesting as such since it enables us to identify the determinant factors at an early point in the chain of ‘noticing, reading, and citing’ pieces of research.

Explanatory variables

The following will test a wide range of explanatory variables. Some of those variables consider themes that research on journal article citations has found to be significant while others are related to specific features of SSRN papers. All data were collected from the SSRN abstract page as it appeared at the moment of the data collection. Accordingly, the variables do not cover information that is not included on the abstract page, such as the academic position (seniority) of the authors. It was also not possible to

⁴⁵ See <http://ssrnblog.com/2009/08/25/ssrn-development-citereader/> and www.ssrn.com/update/all/CiteReader.html.

⁴⁶ Similar Chris Armbruster, ‘Whose Metrics? Citation, Usage and Access Metrics as Scholarly Information Service’ (2010) 23 *Learned Publishing* 33, 36.

check whether and when the abstract page may have changed within the two years, for example, as far as some authors may have moved universities.

Table 2: Descriptive statistics of main explanatory variables (n: 1107)

	Mean	Median	Std. Deviation	Minimum	Maximum
Words in Title	11.49	10	6.476	1	78
Words in Abstract	199.24	173	150.384	9	2422
Pages	32.47	28	25.995	1	378
Accepted Series	0.7	1	0.458	0	1
JEL Non-Law	0.23	0	0.422	0	1
Corporate Topic	0.1518	0	0.35895	0	1
IP Topic	0.0822	0	0.2748	0	1
Constitutional Topic	0.187	0	0.39008	0	1
International Topic	0.1734	0	0.3788	0	1
Empirical Topic	0.1355	0	0.34241	0	1
New Topic	0.3866	0	0.4872	0	1
Review Topic	0.1337	0	0.34048	0	1
Foreign Language	0.0307	0	0.17262	0	1
Top-20 University	0.15	0	0.356	0	1
US Author	0.5881	1	0.4924	0	1
Number of Authors	1.41	1	0.895	1	10
January / October	0.5	0	0.5	0	1
Days online	740.85	741	3.677	735	747
Multi-paper authors	14 dummy variables for authors with five or more papers in either period				

Table 2 summarises the explanatory variables and presents basic descriptive statistics. The rationales for including these variables are as follows:

Citation-based research has found that articles with shorter titles and more pages are cited more often.⁴⁷ In the present case, it is also appropriate to test the length of SSRN paper titles, not least since one of the most downloaded papers has a potentially controversial, one word title.⁴⁸ Furthermore the length of the abstract and the entire paper have been examined: here too, a catchy text may be more appealing; alternatively, it is conceivable that readers may prefer a longer and more informative text.

The next variable reflects that, in the SSRN terminology, papers can be part of either the ‘accepted paper’ or the ‘working paper’ series. The former refers to every publica-

⁴⁷ Eg, Mingers and Xu (n 12); Ayres and Vars (n 14).

⁴⁸ Fairman (n 39). It was also checked whether the number of letters per word (ie short words being preferable) may matter – however, this was not found to be the case in any of the specifications.

tion that ‘has appeared, or [is] selected to appear, in a Paper Series, Journal or Book’.⁴⁹ This can provide an esteem indicator that fosters abstract views and downloads.

In terms of subject matter, SSRN asks authors to classify their papers according to the categories of the Journal of Economic Literature (JEL).⁵⁰ In the present case, it was identified whether authors indicated JEL codes of non-legal topics. It would also be interesting to scrutinise distinctions within the JEL’s law category ‘K’. Yet, the problem is that the authors of about half of the papers of the sample have not provided information about the JEL codes.⁵¹ Thus, in the present study, a number of keywords were identified in order to identify areas of law and subsequently it was checked whether any of these words were mentioned in the abstract of each paper. This approach coded information for corporate law, intellectual property law (IP law), constitutional law and international law (both widely understood).⁵² In addition, keywords were used to identify the potential effect of research on empirical topics, the attractiveness of review papers, and research that claims to present innovative ideas.⁵³

Another factor that is likely to matter is the language of a paper. While the majority of SSRN papers are in English, there are also some papers published in other languages (to be precise: 3.07% in the sample, see Table 2). These latter papers may potentially have fewer downloads and abstract views due to their smaller market.

Next, three author-related characteristics were considered: first, it was identified whether one of the authors belongs to a university of the top-20 in the Times Higher Education World University Rankings 2013-14 – similar to research that found that authors from top institutions receive more citations.⁵⁴ Second, it was coded whether at least one of the authors is at a US institution. This reflects research that found, in terms

⁴⁹ SSRN, FAQs (2014), available at www.ssrn.com/update/general/ssrn_faq.html. Those SSRN paper may either be the final versions or earlier drafts.

⁵⁰ See www.aeaweb.org/econlit/jelCodes.php.

⁵¹ Possibly, because the JEL codes may be unfamiliar to many legal scholar (or, else, that they are seen as unsuitable for many legal researchers due their origins in economics).

⁵² Search terms: (i) ‘corporate’, ‘corporation’, ‘company’, ‘companies’, ‘director’, ‘shareholder’, ‘manager’; (ii) ‘intellectual property’, ‘patent’, ‘copyright’, ‘trademark’; (iii) ‘constitution’, ‘human right’; and (iv) ‘international’, ‘transnational’, ‘comparative’.

⁵³ Search terms: (i) ‘empirical’, ‘statistic’, ‘regression’, ‘data’, ‘quantitative’; (ii) ‘review’, ‘overview’; (iii) ‘new’, ‘innovative’, ‘interesting’, ‘important’, ‘novel’, ‘modern’.

⁵⁴ Eg, Mingers and Xu (n 12).

of SSRN citations, the US being ahead in the ‘author-level Eigenfactor scores’.⁵⁵ As 58.81% of the SSRN papers are (co-) written by US authors,⁵⁶ it also seems plausible to test whether those authors have a larger readership – or else, a possible bias of SSRN for US legal research. Third, there seems to be a shift away from single-authorship in legal publications.⁵⁷ In the current context, co-authorship may help the dissemination of one’s research and the attention it receives. Thus, the subsequent regressions include a variable on the number of authors of each paper.⁵⁸

The final variables reflect the time-dimension indicated in the previous section. The first two of them control for potential differences between the January and October data as well as the slight differences in the number of days each paper has been online. The corresponding variable for January/October has only a very weak correlation with each of the other variables.⁵⁹ Thus, while January papers do not seem to be fundamentally different from the October ones, this variable is necessary since the total number of downloads and abstract views has slightly decreased from January to October (see Table 1). The ultimate variable considers that 14 authors (or identical groups of authors) have uploaded five or more papers in either the January or the October period.⁶⁰ This block uploading may lead to unusual download patterns which may either be negative (say, due to an oversupply of papers by the same author on related topics) or positive (say, if papers by the same author crosslink to each other and therefore stimulate downloads). The variables for these authors aim to control for these factors.

⁵⁵ West et al (n 35) 794. This score aims to indicate the authors influence as a network-based measure.

⁵⁶ See Table 2. For papers (co-) authored by scholars from the UK, Canada, and Australia these figures are 7.86%, 4.61% and 3.52%.

⁵⁷ See also Tom Ginsburg and Thomas J Miles, ‘Empiricism and the Rising Incidence of Co-authorship in Law’ (2011) *University of Illinois Law Review* 1785.

⁵⁸ This was capped at ten due to one outlier in the sample (a report with 46 authors).

⁵⁹ The mean of the absolute values of the correlations is 0.031 and the median is 0.025.

⁶⁰ The precise numbers are: 5-5 (ie five authors with five papers), 1-6, 2-7, 2-8, 1-9, 1-10, 2-15.

Regression model and results

This article examines the determinants for three dependent variables: abstract views, paper downloads, and paper downloads per abstract views.⁶¹ The corresponding regression models depend on the nature and shape of these dependent variables as well as their respective error terms.⁶² The abstract views and downloads are count data. This indicates a Poisson or negative binomial distribution with a Generalized Linear Model (GLM). In the present case, negative binomial is preferred due to overdispersion. The downloads per abstract views are continuous but do not have a normal distribution. Thus, in this model, the subsequent regression uses a Generalized Linear Model (GLM) with Gamma-distributed dependent variables (with log link), which is similar to the negative binomial regression model but considers the continuous nature of those data.

⁶¹ See Table 1, above and corresponding text.

⁶² The following text aims to provide a brief technical summary. For further technical details see, eg, Rainer Winkelmann, *Econometric Analysis of Count Data* (Berlin, Springer 5th ed 2008); Stefany Coxe, Stephen G West, and Leona S Aiken, 'The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives' (2009) 91 *Journal of Personality Assessment* 121.

Table 3: Regression results of SSRN indicators

Explanatory variables: ⁶³	Dependent variables					
	Abstract views (GLM negative binomial)		Paper downloads (GLM negative binomial)		Downloads per views (GLM gamma with log link)	
Words in Title	-.012***	.002	-.020***	.000	-.007***	.003
Words in Abstract	.001***	.001	.001*	.066	.000	.194
Pages	.003**	.012	.003	.152	.000	.819
Accepted Series	-.033	.605	-.116	.182	-.059**	.039
JEL Non-Law	.106	.150	.007	.945	-.105***	.003
Corporate Topic	.258***	.001	.467***	.000	.166***	.000
IP Topic	.271**	.014	.161	.246	-.052	.160
Constitutional Topic	-.004	.954	-.025	.810	-.055	.119
International Topic	.115*	.095	.223**	.016	.124***	.000
Empirical Topic	.133	.125	.096	.435	-.002	.959
New Topic	.000	.995	.041	.615	.013	.638
Review Topic	-.152**	.033	-.167*	.086	-.032	.394
Foreign Language	-.273**	.022	-.362*	.053	-.051	.588
Top-20 University	.339***	.000	.402***	.000	.097***	.002
US Author	.330***	.000	.289***	.001	-.055*	.050
Number of Authors	.037	.285	.043	.368	.003	.860
January / October	.236***	.000	.176**	.028	-	-
Days online	-.003	.702	-.002	.833	-	-

Significance at: *** 1% level; ** 5% level; * 10% level

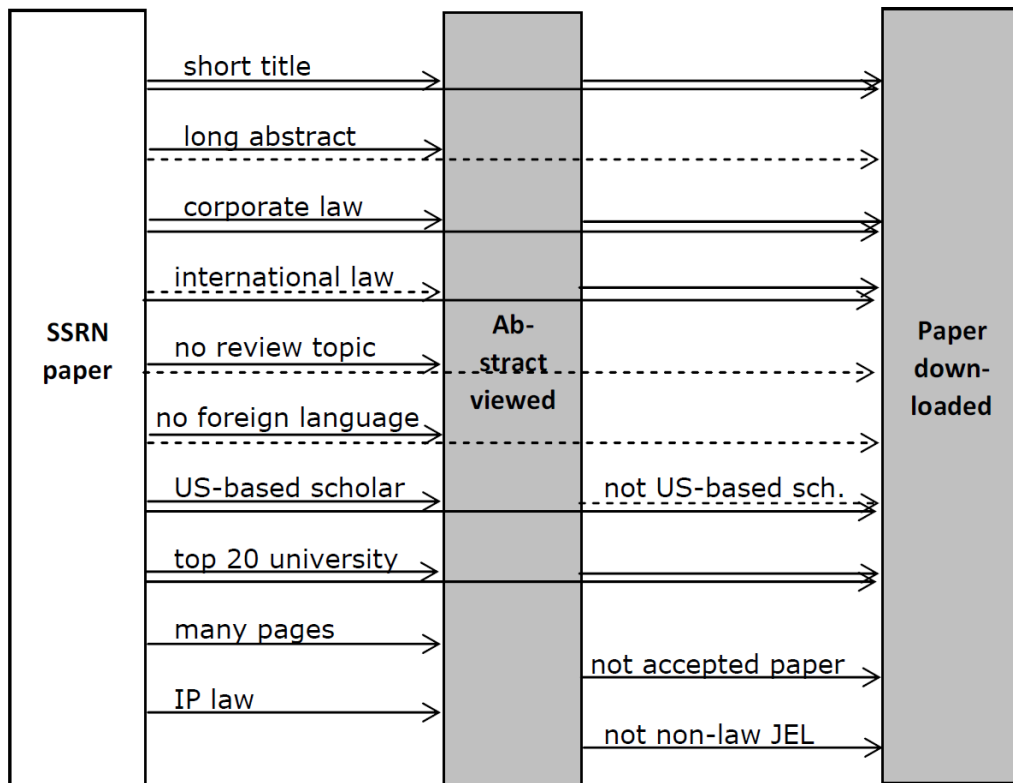
Table 3 reports the coefficients and significance levels of the regression results (for further interpretation see the next section). For GLM regressions it is not possible to calculate information on the ‘ R^2 ’ which in other types of regressions indicates how much variation is explained by the model in question. While there are some suggestions to calculate the predictive power of a model in GLM regressions, it has been said that ‘most of these have serious limitations’ and that none has achieved ‘strong acceptance’.⁶⁴

The three regression models display similar but not identical results. To illustrate the relationship between the models, Figure 2 displays the variables that have a significant effect, with dotted lines used if the significance is only at the 10% level.

⁶³ This table omits the results for the 14 dummy variables that control for multi-paper authors (see Table 2, above).

⁶⁴ Beiyao Zheng and Alan Agresti, ‘Summarizing the Predictive Power of a Generalized Linear Model’ (2000) 19 *Statistics in Medicine* 1771.

Figure 2: Visual presentation of main results



The results of the first two models show a common pattern: it is good to have a short title, a long abstract, to publish on a corporate and international topic, to be at a top-20 university and from the US, and not to publish a review paper or in a foreign language. For abstract views it also helps to have a long paper and to write about IP law. Some of the results of the third model are similar: it is good to have a short title, to write about corporate and international law and to be from a top-20 university. It may seem more puzzling why here being in an accepted series, a non-law JEL code and US authorship has a significant negative effect. The following section tries to make sense of all of those results.

Interpretation and discussion

The signs of the coefficients in Table 3, above, indicate whether the variable in question has a positive or negative effect. Yet, apart from that, the coefficients of GLM regressions do not lend themselves to intuitive interpretation as easily as other regression

models. Thus, to be able to compare, it is helpful to calculate the percentage impact of a one standard deviation increase.⁶⁵ This is reported in Table 4, with the shades indicating the significant coefficients of Table 3. These results will be discussed in the following.

Table 4: Interpretation of coefficients – change per one standard deviation

	Abstract views	Paper downloads	Downloads per views
Words in Title	-7.72%	-12.82%	-4.52%
Words in Abstract	15.05%	15.05%	0.00%
Pages	7.81%	7.81%	0.00%
Accepted Series	-1.49%	-5.02%	-2.62%
JEL Non-Law	4.72%	0.30%	-4.21%
Corporate Topic	10.57%	21.36%	6.48%
IP Topic	8.55%	4.80%	-1.39%
Constitutional Topic	-0.16%	-0.96%	-2.09%
International Topic	4.62%	9.46%	5.00%
Empirical Topic	4.87%	3.45%	-0.07%
New Topic	0.00%	2.04%	0.64%
Review Topic	-4.80%	-5.24%	-1.07%
Foreign Language	-4.12%	-5.24%	-0.86%
Top-20 University	14.37%	17.62%	3.63%
US Author	19.25%	16.50%	-2.64%
Number of Authors	3.37%	3.93%	0.27%
January / October	13.31%	9.62%	-
Days online	-1.10%	-0.73%	-

First, the quantitative indicators about the length of the title, the abstract and the entire paper seem to go in different directions. The relevance of a short title in all three models is plausible: given the high number of papers on SSRN, it is crucial to be quickly able to capture the readers' attention. But, then, it is also necessary to convince them about the substance of the paper: thus, according to the regression results, longer abstracts are better than shorter ones. Interestingly, this already affects the abstract views, presumably because many users come across papers via the SSRN email alerts⁶⁶ which already contain the abstracts.

The number of pages, however, is only visible on the abstract page: thus, in this respect, the statistically significant effect on the abstract views may be more puzzling.

⁶⁵ Cf Coxe et al (n 62) 124.

⁶⁶ Called 'eJournals', see www.ssrn.com/en/index.cfm/lsn/lsn-ejournals/. The default option is that all SSRN papers are included in these email alerts. For the exceptions see www.ssrn.com/en/index.cfm/ssrn-faq/#distribution_eligibility.

However, it would also be implausible to suggest that, prior to downloading a paper, readers crucially consider the number of pages – and, indeed, there is no such effect in the third model. It therefore seems more likely that the length of a paper reflects something else, namely that relatively long papers provide readers with a more extensive and diligent treatment of a particular issue, and that therefore these papers get wider acceptance.⁶⁷

Second, the variables ‘accepted series’ and ‘JEL non-law’ are only significant in the third model. Here, the negative effect of ‘accepted series’ (also in the first two models) is likely to be due to the fact that readers of papers published in journals may be keen to get the published final version of the paper, eg, to be able to cite the page numbers properly. With respect to JEL non-law papers which are part of the SSRN LSN series, the general expectation may be that these interdisciplinary papers have a larger audience. This may therefore explain the (narrowly not statistically significant) result of the first model that JEL non-law papers have increased abstract views. However, as this readership is more diverse than for other papers, it can then also happen more often that the abstract does not lead to an actual download of the paper: thus, in this respect, there is the plausible negative effect in the third model.

Third, the results in the subject-matter categories show that writing about corporate law, international law, and (to a lesser extent) IP law has a positive effect. SSRN is very popular in corporate legal scholarship since many of those legal topics also interest more business-oriented researchers and practitioners across countries.⁶⁸ Similarly, many topics of IP law have an international and interdisciplinary dimension as well as an extra-academic audience. This can be contrasted with constitutional law, an area of law that has traditionally been narrower, ie oriented towards legal doctrine and domestic law.⁶⁹

⁶⁷ Such reasoning can be corroborated by the finding that in the UK’s Research Assessment Exercise 2008 the length of articles was found to be a significant determinant for the assessment of research outputs. See Mathias Siems, ‘What explains the RAE 2008 law output results?’ (3 May 2013), available at <http://siemslegal.blogspot.co.uk/2013/05/what-explains-results-for-rae-2008-law.html>.

⁶⁸ Eg, there is a separate Corporate Governance Network on SSRN, available at www.ssrn.com/en/index.cfm/cgn/. The European Corporate Governance Institute’s (ECGI) also has its own Law Working Paper Series on SSRN: http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=291302.

⁶⁹ Brian R Cheffins, ‘The Trajectory of (Corporate Law) Scholarship’ (2004) 63 *Cambridge Law Journal* 456, 505 (difference between corporate and constitutional scholarship).

For topics of international law, the wider international appeal of such papers is a plausible reason why those are frequently viewed and downloaded. It can also be suggested that the popularity of the category ‘international law’ (which includes comparative and transnational law)⁷⁰ has a qualitative dimension since ‘deep legal scholarship’ is often said to require that the researcher needs to go beyond the solutions provided by the researcher’s own legal system.⁷¹ Such a line of reasoning may also be applied to the positive (though not significant – narrowly in the first model) effect of empirical legal research. By contrast, the negative (and significant) effect of review topics may show that the SSRN audience values original scholarship⁷² more than a summary of the state of the art.

Fourth, the results in the categories language and US authorship point towards similar but not identical interpretations. Since the main language of SSRN is English, anyone who does not read (or does not want to read) materials in English may not actively search on the SSRN website or subscribe to their email alerts. As many SSRN users are from the US, it can also be a benefit to be a US author since there is likely to be a greater overlap between the subject matter interest of those authors and US readers than between non-US authors and US readers. Alternatively, the positive effect of being a US author can be related to the role of name recognition because the names of US legal scholars and/or their universities are more likely to be familiar to US readers. The negative effect in terms downloads per views is not inconsistent with those explanations since the greater availability of US materials can mean that readers that have accessed the abstract page of a paper are then more selective prior to the download of the paper.

Fifth, the popularity of papers by scholars from top-20 universities is very strong, but the interpretation of this effect is not straight-forward. The most intuitive reason is that readers know and value the name of those top universities. It is also plausible that professors at these universities tend to be relatively well-known: thus, in addition, the name recognition could be related to the author in question. Moreover, and more contentiously, it may be argued that these authors are just ‘better’ in the way they write their papers

⁷⁰ See n 52.

⁷¹ Cf Jan M Smits, *The Mind and Method of the Legal Academic* (Cheltenham: Edward Elgar, 2012).

⁷² To be sure, identifying ‘originality’ is not always straight-forward. See Mathias Siems, ‘Legal Originality’ (2008) 28 *Oxford Journal of Legal Studies* 147.

and how they formulate their titles and abstracts; thus, the relevance of the top-20 universities may also capture the quality of their SSRN submission.

Wider implications for SSRN users?

It may be asked whether the findings of this article can be read as telling SSRN users ‘how to increase their SSRN downloads’, for example, have short titles but long abstracts, publish on corporate and international law topic, and get an affiliation with a top-20 university. In their corresponding discussion of citation metrics, Ian Ayres and Fredrick Vars offer a cautionary note:

‘There may be a strong temptation to read many of our results as recipes for citation success. Authors (or law review editors) might think that they could increase their citations if they just publish longer articles or shift toward publishing constitutional law pieces. Such inferences are fraught with peril. The fallacy of aggregation suggests that just because long articles have tended to be cited more in the past does not mean that journals should force authors to add 10 pages of pabulum to their articles in order to generate more citations.’⁷³

In the present case, there are also good reasons to be wary since it cannot be excluded that some of the explanatory variables proxy for the unobservable quality of a paper. This line of reasoning was specifically mentioned for the length of a paper, writings on international law, and the effect of being at a top-20 university, but it may also matter for some of the other variables (eg, a short title may be due to the clear focus of the paper; a long abstract may be due to the substantive contribution of the paper).

⁷³ Ayres and Vars (n 14) 446-7.

CONCLUSION

The theoretical section of this article suggested a distinction between internal and external factors that may contribute to the reception of scholarship.⁷⁴ Based on a sample of 1107 papers of SSRN's Legal Scholarship Network, the empirical findings of this article show that both sets of factors indeed play a role: on the one hand, it pays off to have a catchy title but a relatively long abstract and paper; on the other hand, US authorship and a top-20 university affiliation were also found to be significant. Writing about corporate, IP and international law also tends to get more attention on SSRN – which an author with no expertise in these areas may perceive as an external factor.

These findings had the aim to fill a gap in the literature which has largely been concerned with the determinants of journal citations only.⁷⁵ A controversial issue of the current discussion is how much importance one shall give to any particular bibliometric information.⁷⁶ It is interesting to note that Thomson Reuters, one of the publishers involved in bibliometrics, admits that:

‘No one metric can fully capture the complex contributions scholars make to their disciplines, and many forms of scholarly achievement should be considered’.⁷⁷

A recent initiative called ‘Altmetrics’ also aims to broaden the scope of information that should be taken into consideration, namely: (i) usage, ie downloads and view; (ii) peer-review, ie expert opinion; (iii) citations; and (iv) altmetrics in a narrow sense, ie storage, links, bookmarks, and conversations.⁷⁸

This suggestion of a combined approach is also appropriate in the current context. While, to some extent, SSRN downloads and abstract views may capture the unobservable quality of a paper,⁷⁹ such information cannot replace peer-review assessments since – according to the findings of this article – factors not related to quality also play a role

⁷⁴ See text to n 5.

⁷⁵ See references in n 12-15.

⁷⁶ See, eg, the references in n 18, 24, 34.

⁷⁷ ‘Statement Regarding the San Francisco Declaration on Research Assessment’, available at http://researchanalytics.thomsonreuters.com/statement_re_sfdra/ (though the subsequent text then also tries to justify the various indicators used by Thomson Reuters).

⁷⁸ ‘Altmetrics: A Manifesto’, available at <http://altmetrics.org/manifesto>. See also José Luis Ortega, ‘Relationship between Altmetric and Bibliometric Indicators Across Academic Social Sites: The Case of CSIC's Members’, (2015) 9 *Journal of Informetrics* 39.

⁷⁹ See text to n 73.

for the attention it receives. Such a cautionary approach is in line with the corresponding view that journal impact factors should not be used to assess the quality of individual pieces of research.⁸⁰

In order to fully understand the relationship between SSRN indicators and the quality of legal scholarship, it would also be necessary to have quantifiable data on the quality of the research outputs in question. Related research has been conducted for the relationship between other indicators in other disciplines.⁸¹ Future research may therefore ask a group of expert reviewers to do a ‘mock research assessment’ of a random sample of SSRN law papers in order to establish how far a peer-review assessment may (or may not) be correlated with data on downloads and abstract views.

⁸⁰ See n 24.

⁸¹ Elizabeth S Vieira, José AS Cabral, and José ANF Gomes ‘How Good is a Model Based on Bibliometric Indicators in Predicting the Final Decisions Made by Peers?’ (2014) 8 *Journal of Informetrics* 390; Ludo Waltman, Nees Jan van Eck, and Thed N van Leeuwen, Martijn S Visser, Anthony FJ van Raan, ‘On the Correlation between Bibliometric Indicators and Peer Review: Reply to Opthof and Leydesdorff’ (2011) 88 *Scientometrics* 1017. See also n 18.