

## Research Dissemination and Impact: Evidence from Web Site Downloads

LEE PINKOWITZ\*

### ABSTRACT

The *Journal of Finance* Web site disseminates research expediently to a broad audience. Papers were downloaded 284,170 times from November 1997 to November 1999. The average paper receives 85 downloads per month and is available 10 months before publication. Articles are downloaded more than shorter papers. Lead articles are downloaded more than other articles because they are of greater interest, and they receive an endorsement as the lead. Downloads are positively correlated with citations and may be a useful measure of research impact. Finally, placing forthcoming articles online does not adversely affect subscriptions and may increase the SSCI impact factor.

THE PROLIFERATION OF THE WORLD WIDE WEB (hereafter, Web) has dramatically changed the way information is transmitted. For academic journals, the Web provides an opportunity to make articles available to subscribers and the public while eliminating the delay that is inevitable with a print publication. The *Journal of Finance* created a Web site in 1995 and began posting full-text versions of forthcoming papers in early 1996. After several years, the question remains, has the Web site been successful at distributing research in a timely and effective manner?

This paper analyzes activity on the *Journal of Finance* Web site, specifically the number of papers downloaded. Examining downloads provides insight into the dissemination of research that was previously unavailable. Assuming a download corresponds to someone reading the paper, we can, for the first time, quantify which articles generate the most interest.

This information is likely to be important to three separate constituencies. First, editors would be able to determine the most popular papers, which can help them shape the focus of the journal to match readers' interests. Second, because research productivity is a major factor in the tenure decision at most universities, downloads may be valuable to authors and university administrators. Often, universities prefer quality research to mere quantity; however, it is difficult to measure quality in a standardized and systematic

\* Pinkowitz is at the McDonough School of Business, Georgetown University. Kirsten Anderson, Ken Cavalluzzo, Sandeep Dahiya, Allan Eberhart, Jan Jindra, Denise Pinkowitz, Rohan Williamson, and especially an anonymous referee provided helpful comments. René Anvelt and Mo Chen provided capable research assistance. The author received research support from the Capital Markets Research Center at Georgetown University. Any errors are the responsibility of the author.

fashion. Many researchers have argued that citation analysis offers a way to measure both quality and impact (see, e.g., Borokhovich, Bricker, and Simkins (2000)). In this paper, I claim that downloads may be an additional way to measure research impact. Finally, I examine whether making forthcoming articles available online affects subscriptions and citation impact factors, which may interest other journals debating whether to make articles available on the Web.

In the 25 months from November 1997 to November 1999, the *Journal of Finance* posted 202 articles and shorter papers for a total of 3,357 paper months. The papers were downloaded more than 284,000 times, averaging nearly 85 per paper, per month. Additionally, articles were available an average of 311 (median of 320) days prior to publication. It appears the Web site allowed for research to be disseminated in a more expedient manner than has traditionally been the case.

I find that a paper's general topic area is associated with its number of downloads. Papers categorized as "General Financial Markets" are downloaded more often than papers in other areas. Papers categorized as "General Finance" are downloaded less frequently, while papers in the areas of "Financial Institutions," "Corporate Finance and Governance," and "Money and Interest Rates" are downloaded with approximately equal frequency.

Borokhovich et al. (2000) show that lead articles in the *Journal of Finance* are cited more frequently than other articles. I find that lead articles are downloaded significantly more often than other papers both before and after being named the lead article. The fact that they are downloaded more often before being designated as the lead suggests they may be more influential. However, there also appears to be certification value to the lead position. Once a paper is assigned as the lead article, the number of downloads significantly increases. The results suggest that an editor can influence what people will read.

Downloads are also positively correlated with future citations. On average, it takes 109 downloads per month to generate one citation per year from another researcher. The results suggest that downloads, like citations, may be a useful measure of research impact. Finally, making forthcoming articles available via the Web does not appear to adversely affect subscriptions and may increase journal impact factors.

The next section explains why downloads may be a beneficial measure of research impact while Section II describes the data. Section III studies the effectiveness of the Web site in disseminating information, analyzes the characteristics of downloaded papers, and examines whether downloads are correlated with citations. Section IV examines whether making forthcoming articles available online affects subscriptions and journal impact factors. Section V concludes.

## I. Downloads versus Citations

There is a broad base of literature examining citation counts of academic research (see, e.g., Borokhovich et al. (2000)). Research impact is a key element in the promotion decisions at most universities; thus, it is crucial to

develop a quantitative measure of impact. Citation analysis claims to do this; however, there are some problems that might be avoided using Web downloads.

One weakness with citation analysis is that it cannot measure impact on practitioners who do not publish research. Because practitioners as well as academics read articles, citation analysis may bias against scholars whose research is focused on real-world applications. Downloads do not suffer from this problem. At the same time, practitioners might be less adept at identifying quality papers; hence, while downloads may measure readership, they might overstate research quality.

Publication lag is potentially another problem with citations. The mean (median) time between when a paper is posted on the *Journal of Finance* Web site and its appearance in print is 311 (320) days. Thus, it takes 10 months, on average, from the time the paper is *accepted* until others see it in print. If the citing author faces a similar delay, it may take a year and a half before citation analysis will benefit the original author. Additionally, this does not account for the time that the citing author will need to develop the idea, write the paper, and have it proceed through the refereeing process. Downloads allow impact to be measured as soon as the paper is posted online.<sup>1</sup>

Finally, citation analysis relies on the ability of other authors to be published. Arguably, the goal of research is to have it read by others; thus, a measure of impact should be the breadth with which ideas are disseminated. Using citations likely understates the impact of research. For a citation to appear in the Social Science Citation Index (SSCI), an author must not only read the article, but then must publish a related paper. However, downloads can capture scholars who read the paper even if it is outside their research area and thus never cited in any of their work.<sup>2</sup>

## II. Data

The data are from the log files of the *Journal of Finance* Web server housed at The Ohio State University. The logs record all requests for files and allow for a quantitative study of how many times Web pages are accessed. The *Journal of Finance* Web site started in 1995 and began posting full text versions of accepted papers in early 1996. Through December 1999, 268 articles and shorter papers were posted. This study includes access data for

<sup>1</sup> Of course, working papers are cited, which could mitigate the problems with publishing lag. However, when I examine citations of 129 papers appearing in the December 1997 to October 1999 issues of the *Journal of Finance* (see Section III), I find that 75 percent (median 83 percent) of the total cites are to the *Journal of Finance* version and not the working paper. Thus, it does not appear that citations to working papers solve the problem of publication lag.

<sup>2</sup> Downloads may overstate impact since it is inevitable that some people will not read each paper they download. In addition, some people may download an article more than once. However, a similar argument can be made about citation counts because authors may cite the same paper in several of their own papers. Also, some authors may cite papers that they have never actually read.

the 25 months from November 1997 to November 1999. Further, I use the SSCI database to gather citation and journal impact factor data.

For each paper, I record whether it appeared in the articles or shorter papers section of the *Journal of Finance*. In addition, I record the date the article was made available to the public. For papers made available on or after the 20th of the month, I assume the paper was posted the next month. This prevents papers available for only a few days from being counted as accessible for a full month.

For papers published prior to December 1997, the online version was removed from the Web site once the print version of the *Journal of Finance* was sent to subscribers. Thus, I eliminate all articles which appeared in print prior to December 1997. I remove Hayne Leland's Presidential Address, published in the 1998 Proceedings issue, since the presidential address is always the lead article. Thus, it does not allow for inferences regarding which article the editor selects as the lead.<sup>3</sup> Finally, I eliminate any paper posted after October 20, 1999, so there is at least one month of download data for each paper. The resulting sample includes 202 papers composed of 115 articles and 87 shorter papers.

Each month, I examine the number of downloads for each paper using the Web server log files. I count a download only the first time it is requested from a specific IP address in any given month. This mitigates the problem of double-counting multiple downloads from the same reader.

### III. An Examination of Web Site Downloads

Table I shows summary statistics by month. The first column shows the number of papers that were available on the Web site, which ranges from 56 in November 1997 to the full sample of 202 in November 1999. Over the two-year period, a total of 284,170 downloads were recorded, or about 85 downloads per paper, per month.

Table II shows descriptive statistics segmented by subject. The subject areas are the same the *Journal of Finance* uses for the Annual Reports in the Proceedings issue. "General Financial Markets" has the majority of papers (122 out of 202), and the most downloads, with an average of 88.84 per paper, per month. "Corporate Finance and Governance" has the second most papers (49 out of 202) and the second highest number of downloads (83.39). The medians have a similar ordering except for "Money and Interest Rates," which has the highest median but only the third highest mean.

The positive correlation between the number of papers in each category and the number of downloads each paper in that category receives may indicate a degree of specialization.<sup>4</sup> It seems that the *Journal of Finance* is a

<sup>3</sup> This is the only Proceedings paper and the results are similar when I include it in the sample.

<sup>4</sup> The correlation coefficient between the number of papers in a category and mean downloads is 0.7237 with a *p*-value of 0.1669. However, the small sample of five observations makes finding significance difficult.

**Table I**  
**Summary Statistics by Month**

Descriptive statistics are shown for downloads from the *Journal of Finance* Web site for each month in the sample. Papers Posted is the number of papers available on the Web site during that month. Mean, Median, Minimum, and Maximum refer to number of downloads per paper.

Month	Papers Posted	Total Downloads	Mean, per paper	Median, per paper	Minimum, per paper	Maximum, per paper
11/97	56	4,561	81.45	66.0	26.0	213.0
12/97	65	5,650	86.92	66.0	22.0	308.0
1/98	68	6,790	99.85	73.5	32.0	453.0
2/98	79	8,387	106.17	84.0	26.0	372.0
3/98	81	8,507	105.03	78.0	19.0	495.0
4/98	92	9,349	101.62	77.5	16.0	440.0
5/98	97	8,542	88.06	58.0	12.0	497.0
6/98	106	9,302	87.76	66.0	10.0	297.0
7/98	116	10,408	89.72	64.0	15.0	316.0
8/98	121	9,456	78.15	51.0	11.0	276.0
9/98	130	11,321	87.09	64.5	12.0	379.0
10/98	132	11,648	88.24	65.5	10.0	375.0
11/98	139	13,658	98.26	76.0	17.0	417.0
12/98	147	13,523	91.99	72.0	13.0	495.0
1/99	152	13,852	91.13	66.5	10.0	408.0
2/99	157	14,348	91.39	75.0	20.0	485.0
3/99	161	15,549	96.58	71.0	16.0	489.0
4/99	162	15,095	93.18	69.0	14.0	956.0
5/99	165	13,875	84.09	61.0	13.0	675.0
6/99	169	12,491	73.91	55.0	12.0	474.0
7/99	177	11,756	66.42	48.0	7.0	504.0
8/99	184	12,672	68.87	48.5	8.0	426.0
9/99	197	14,735	74.80	58.0	15.0	367.0
10/99	202	14,540	71.98	52.0	5.0	369.0
11/99	202	14,155	70.07	54.0	11.0	422.0
Totals	3,357	284,170	84.65	63.0	5.0	956.0

primary outlet for papers that focus on either “General Financial Markets” or “Corporate Finance and Governance.” In addition, it appears that the majority of the *Journal of Finance*’s readers are interested in those areas. Since those categories are broad, the results are probably better interpreted as showing that the increasing number of specialized journals has allowed the *Journal of Finance* to focus its attention on more mainstream topics. Nonetheless, the papers in the other three areas are also well read, averaging from 58 to 80 downloads per month.

Table III examines downloads in a regression framework. The number of papers available changes each month and the pooled time series is likely to be autocorrelated. Thus, I run cross-sectional regressions each month of the form:

$$Download_i = \alpha + \sum_{n=1}^4 \delta_n Sub_{n,i} + \sum_{m=1}^M \beta_m X_{m,i} + \epsilon_i, \quad (1)$$

Table II  
**Summary Statistics by Subject Area**

Descriptive statistics are shown for downloads from the *Journal of Finance* Web site for each broad subject area in the sample. Papers is the number of papers available on the Web site, while Paper Months is the sum of the number of papers multiplied by the number of months that each paper is posted. Mean, Median, Minimum, and Maximum refer to number of downloads per paper.

Subject Area	Papers	Paper Months	Total Downloads	Mean, per paper	Median, per paper	Minimum, per paper	Maximum, per paper
General Finance	7	153	8,993	58.77	50.0	13.0	213.0
General Financial Markets	122	2,001	177,768	88.84	65.0	5.0	956.0
Financial Institutions	18	262	19,439	74.19	61.0	9.0	296.0
Corporate Finance and Governance	49	823	68,634	83.39	61.0	7.0	495.0
Money and Interest Rates	6	118	9,336	79.12	76.5	24.0	176.0
Total	202	3,357	284,170	84.65	63.0	5.0	956.0

where  $Download_i$  is the number of downloads for paper  $i$  for the month,  $Sub_{n,i}$  are dummy variables representing the subject area of paper  $i$ , and  $X_{m,i}$  are variables which may affect the number of downloads a paper receives. I then use the time series of 25 regression coefficients to make my inferences (see Fama and MacBeth (1973)). The coefficients in Table III can be interpreted as downloads per paper, per month. The adjusted  $R^2$ s are averages from the 25 cross-sectional regressions and provide a measure of the explanatory power of the specification.

The first specification shows that "General Financial Markets" papers are more likely to be downloaded than those in "Financial Institutions" or "General Finance." This result is robust to many specifications and contrasts with that in Borokhovich et al. (2000), which shows that future citations do not depend on subject area. It is possible that a larger percentage of the papers published in the "General Financial Markets" category have less impact than other categories. Thus, it takes more people reading them to generate a citation. On the other hand, it may be that papers in that category appeal to readers that are less likely to publish.

During my sample period, papers were left on the site after they were published, thus I include the variable *Published*, which equals one if the paper has appeared in print, and zero otherwise. As might be expected, the coefficient is negative and significant. Once a paper is published, it receives 23 fewer downloads per month. However, the intercept is around 160, suggesting that even published papers are downloaded fairly often. This may indicate that people use the Web site as a substitute for a print subscription. I examine this question in Section IV.

It may be expected that the longer a paper is available, the fewer downloads per month it will receive. The coefficient on *Months Available* is significantly negative, even with *Published* in the regression, indicating that time decay is present and not solely due to articles appearing in print.

Borokhovich et al. (2000) show that articles are cited more frequently than shorter papers in the *Journal of Finance*. In addition, in his report for 1999 (Stulz, 2000), the editor indicates that shorter papers are characterized less by length than audience size. If this is true, fewer people should download the shorter papers, provided the editor assigns the papers correctly. The significantly positive coefficient on *Article Dummy* indicates this is the case.

Many academics have the impression that a lead article in the *Journal of Finance*, or any journal, indicates a greater contribution to the literature. Schwert (1993) shows that lead articles in the *Journal of Financial Economics* (*JFE*) are cited more frequently than other articles. Borokhovich et al. (2000) confirm this for the *JFE* and find that it holds for the *Journal of Finance* as well.

There are two reasons lead articles may be downloaded more frequently. First, the editor may attempt to assign papers of greater impact/interest as the lead, indicating that papers that appear as the lead article are more influential. Second, people may be more likely to read the lead article because they believe that the editor has certified it to be of greater impor-

Table III  
**Determinants of Downloads**

Downloads are the number of times a paper is downloaded from the *Journal of Finance* Web site. Fama–MacBeth (1973) regressions ((1) to (6)) are shown in the table. For each month, a cross-sectional regression is run using downloads for that month as the dependent variable. The mean of the time series of coefficients is shown in the table. The  $t$ -statistic is based on the mean and the standard deviation of the time series of coefficients. The  $R^2$  is the average from the cross-sectional regressions. Each time series contains 25 months. *General Financial Markets* is the omitted subject dummy variable and is measured by the intercept. *Published* is a dummy variable equal to one if the month is of or after the month that the paper appears in print, and zero otherwise. *Months Available* measures the number of months that the paper has been posted on the site. *Article Dummy* is assigned one if the paper is an article and zero if the paper appears in the shorter papers section. *Lead Article (Shorter)* is a dummy variable equal to one if the paper appears as the lead article (shorter paper), and zero otherwise. The announcement date is the first day of the month of the issue immediately prior to the issue in which the article appears as the lead, until the October 1999 issue. After that, it is the day that the issue was announced online. It measures the date where subscribers would know the order of the issue. The fixed effects regressions ((7) and (8)) include paper-specific dummy variables. Regression eight also includes unreported calendar month dummies to account for variation through time.

	Fama-MacBeth Regressions					Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept (General Financial Markets)	159.07 (23.99)	160.82 (24.61)	161.78 (25.05)	161.78 (25.05)	160.41 (23.14)	160.41 (23.14)		
General Finance	-6.54 (-1.79)	-7.78 (-2.23)	-7.43 (-2.06)	-7.43 (-2.06)	-8.31 (-2.34)	-8.31 (-2.34)		
Financial Institutions	-20.41 (-5.91)	-22.64 (-6.74)	-22.13 (-6.64)	-22.13 (-6.64)	-19.49 (-5.95)	-19.49 (-5.95)		
Corporate Finance and Governance	-0.96 (-0.65)	-7.69 (-4.64)	-7.98 (-4.77)	-7.98 (-4.77)	-0.66 (-0.45)	-0.66 (-0.45)		
Money and Interest Rates	-2.86 (-1.69)	-1.59 (-0.70)	-1.22 (-0.54)	-1.22 (-0.54)	-2.62 (-1.27)	-2.62 (-1.27)		
Published	-23.28 (-3.94)	-22.66 (-3.87)	-23.38 (-3.91)	-23.38 (-3.91)	-23.99 (-3.86)	-23.99 (-3.86)	-31.61 (-13.17)	-35.90 (-14.85)
Months available	-10.55 (-4.85)	-10.52 (-4.92)	-10.74 (-4.90)	-10.74 (-4.90)	-10.84 (-4.66)	-10.84 (-4.66)	-2.48 (-16.45)	
Article dummy	22.70 (16.41)	16.20 (11.70)	16.16 (12.75)	16.16 (12.75)	22.57 (15.39)	22.57 (15.39)		
Lead article		55.54 (13.34)		45.24 (4.79)				
Lead shorter		0.82 (0.48)						
Lead article before announced			45.24 (4.79)					
Lead article after announced			70.08 (11.66)	24.84 (2.09)			16.89 (2.67)	17.39 (2.80)
Lead shorter before announced					-14.40 (-2.45)			
Lead shorter after announced					7.14 (2.87)	21.53 (2.81)	8.51 (1.39)	8.09 (1.35)
<i>N</i>	25	25	25	25	25	25	3,357	3,357
Adjusted <i>R</i> <sup>2</sup>	31.91%	35.16%	36.11%	36.11%	31.24%	31.24%	73.74%	74.92%

tance. Although the Schwert (1993) and Borokhovich et al. (2000) analyses cannot distinguish between these two possibilities, this can be done using the download data.

In specification two, I initially examine whether lead articles are more frequently downloaded. Since the *Journal of Finance* includes both articles and shorter papers, I examine both the lead article and lead shorter paper in each issue. The dummy variable *Lead Article (Shorter)* equals one if the paper is the lead article (shorter paper), and zero otherwise. Controlling for other factors, lead articles are downloaded significantly more frequently, with 55 more downloads per month than other articles. However, the lead shorter paper does not appear to be downloaded more often. Although positive, the coefficient is small and insignificant.

Having shown that lead articles are downloaded more often than other papers, I attempt to determine whether lead articles are downloaded more frequently because they are of greater interest, or if it is the result of editorial certification. When articles are first posted on the *Journal of Finance* Web site, they are listed as either an upcoming article or shorter paper, but no one knows which article (or shorter paper) will be the lead until the paper is assigned to an issue by the editor. Once the paper is assigned, the Web site reflects this by grouping the papers together in the order that they will appear in print. I assume the public learns an article will be the lead from the Miscellanea of the prior issue. Thus, I use the first day of the month of the issue immediately prior to the one in which the article appears. In most cases, the date is two months before the article appears in print. However, beginning with the October 1999 issue, the Web site announced forthcoming issues before the Miscellanea. For those issues, I use the date on which the issue was posted on the Web site.<sup>5</sup>

In specification three, I separate the variable *Lead Article* into two mutually exclusive dummy variables: *Lead Article Before Announced* and *Lead Article After Announced*. The coefficient on *Lead Article Before Announced* is 45.2 and significant, indicating that lead articles are downloaded more often even before readers know the article is the lead. It appears that the editor correctly assigns lead articles.

The coefficient on *Lead Article After Announced* is also positive and highly significant. In fact, it is larger than the coefficient on *Lead Article Before Announced*, indicating that there may also be a certification impact of the lead article. Specification four shows this is the case by displaying the marginal effect of the announcement of the lead article. The coefficient is significant and indicates that downloads increase by 25 per month after the paper is certified as the lead article. Thus, the announcement increases the number of downloads of the article by 60 percent.

<sup>5</sup> For robustness, I also use the date the issue directory was created on the Web server. Even for issues prior to October 1999, this date was, at times, earlier than the Miscellanea. With the earlier dates, I may be less likely to capture the announcement effect, yet the results are similar. Lead articles are read more often prior to being announced, and the announcement effect is positive, but only marginally significant.

Regressions five and six examine whether the same holds true for the lead shorter paper. There is still a certification role, since the marginal effect of the selection is positive and significant. However, the coefficient on *Lead Shorter Before Announced* is negative and significant indicating that the papers assigned to the lead shorter paper position are not downloaded more prior to the announcement. If the editor is attempting to place papers of interest/impact in the lead shorter paper position, the evidence in specifications five and six indicates that he is incorrectly selecting those papers.

To further test whether certification affects downloads, in specifications seven and eight, I run fixed effect regressions using the pooled sample. The regressions include paper-specific dummy variables and are designed to cleanly measure the announcement effect of being assigned as the lead article (shorter paper). Regression seven includes a trend variable for number of months available, while regression eight includes unreported calendar month dummy variables to allow for nonlinear variation. In both cases, the variable *Lead Article After Announced* is positive and significant with a coefficient around 17. Therefore, in the month a paper is announced as the lead, the number of downloads significantly increases. This lends strong support to the idea that the editor's decision acts as certification.<sup>6</sup> The insignificant coefficients on *Lead Shorter After Announced* indicate that the certification role is weak for shorter papers.

Another issue is whether academics and administrators can rely on downloads as a measure of research impact. I examine this question using citation counts for 129 papers published in the *Journal of Finance* from December 1997 to October 1999. The citation data from the SSCI database are current as of November 2000, giving me at least one year after publication to allow for cites to be recorded.

To measure research impact, I eliminate self-citations and only examine the number of cites a paper receives from outside authors. This should give the best indication of whether downloads correspond to citations, as authors should not need to download their own papers.

To examine whether downloads are correlated with citations, I estimate the following regression:

$$\begin{aligned} \text{Cites per year} = & 0.647 + 1.933 \text{ Article} + 0.009 \text{ Downloads} & N = 129 \\ & (1.48) \quad (4.20) \quad (2.25) & \text{Adj. } R^2 = 0.1613 \end{aligned}$$

where *Cites per year* is the number of citations per year from other authors as of November 2000, *Article* is a dummy variable equaling one if the paper was in the articles section of the *Journal of Finance*, and *t*-statistics are in parentheses.<sup>7</sup> *Downloads* is the average number of downloads per month, for

<sup>6</sup> The results in regressions seven and eight are conservative, because if I have mismeasured the month when readers become aware of which article will be the lead, the coefficient on *Lead Article After Announced* should be biased toward zero.

<sup>7</sup> Cites per year is calculated by dividing total citations by (months between November 2000 and the month published)/12. Thus, if an article was published in the December 1997 issue, its citations are divided by (35/12) because the paper was published for 35 months as of November 2000.

months prior to publication.<sup>8</sup> Consistent with Borokhovich et al. (2000), articles are cited more frequently than shorter papers. Additionally, downloads are significantly related to future cites. The coefficient implies that a paper receiving 109 downloads per month will generate one citation per year.

I examine the robustness of the result in several ways. First, I use different measures of citation count as the dependent variable. I examine the results including self-citations, as well as restricting the analysis only to cites of the paper in the *Journal of Finance* rather than a working paper version. I also analyze the results if I use download data for all months (November 1997 to November 1999), rather than just months prior to publication. Finally, I use cumulative downloads rather than average downloads per month. In each case, the results are similar to those reported.

I examine other specifications with additional control variables. My results are similar if I include dummy variables for the subject area of the paper and/or if I include a dummy variable measuring whether the paper was the lead article. Since 19 of the 129 papers have zero citations as of November 2000, I reexamine my results using Tobit regressions and find nearly identical results.

Overall, the results indicate that downloads appear to foreshadow future citations. Thus, downloads may be another useful measure of research impact. However, the estimate of 109 downloads per month should be used with caution. In the robustness checks, the implied number of downloads per month needed to generate one citation per year ranged from 63 to 150.

#### **IV. Effect of Posting Forthcoming Articles on Subscriptions and Impact Factors**

Although forthcoming articles appear to be frequently downloaded, journals may need to consider the impact that the practice may have on its subscriber base. If readers can download articles in the journal at no cost, this may lower the incentive to pay money for a print subscription. At the same time, making articles available online to nonsubscribers may increase citations to those papers, which could boost a journal's SSCI impact factor (see Borokhovich et al. (2000) for a description). Thus, there may be both costs and benefits to making articles available online. This section examines whether subscriptions and impact factors are affected by making forthcoming articles available online.<sup>9</sup>

<sup>8</sup> The mean (median) number of months papers are available prior to publication is 8.05 (9.00), with a range of 1 to 16. Roughly 75 percent of the papers have at least five months of download data prior to publication. Of the 129 papers, the mean (median) number of citations is 5.63 (3.00), with a range of 0 to 28.

<sup>9</sup> I am grateful to Marcia Cornett, David Pyle, and Jack Rader for providing me with subscription data. Additionally, I thank Betty Simkins for sharing her data on impact factors and Martha Auvil and Susan Huot for providing information regarding the *Journal of Financial and Quantitative Analysis*.

**Table IV**  
**Impact of Web Site on Subscriptions and Impact Factors**

Panel A shows the data used in the regressions in Panel B. The dependent variable is the level of annual subscriptions or Social Science Citation Index (SSCI) impact factors. *Impact Factors* measure how often the average article published in a particular journal during the prior two years is cited, and is a proxy for journal quality. *JF Web Site* (*JFQA Web Site*) equals one for the *Journal of Finance* (*Journal of Financial and Quantitative Analysis*) after 1995 (1995), and zero for all other observations. *Association Affiliation* is a dummy variable which equals one if the journal subscription fee is included in the dues of an association. *Cost Per Page* is the annual subscription fee (in 1994 dollars) divided by the total number of pages published in the journal that year. In the Impact Factor regressions, the *Web Site* variables equal one after 1996 because there is a delay of one year before published papers affect impact factors. Missing data is filled using the average of the prior and next year. Both regressions are run with journal-specific and annual dummy variables and have *t*-statistics adjusted for heteroskedasticity.

Journal	Subscriptions	Impact Factors
Panel A: Data Availability		
<i>American Economic Review</i>	1970–1999	1977–1999
<i>Financial Management</i>	1977–1999	1977–1999
<i>Journal of Business</i>	1970–1999	1977–1999
<i>Journal of Finance</i>	1970–1999	1977–1999
<i>Journal of Financial and Quantitative Analysis</i>	1970–1999	1977–1999
<i>Journal of Financial Economics</i>	n.a. <sup>a</sup>	1978–1999
<i>Journal of Political Economy</i>	1970–1999	1977–1999
<i>Quarterly Journal of Economics</i>	1976–1985, 1987–1999	1977–1999
<i>Review of Financial Studies</i>	1988–1990, 1992–1999	1992–1999
Panel B: Regression Analysis		
Association affiliation	-149.13 (-0.48)	
Cost per page	-5,159.51 (-0.77)	
<i>JF</i> Web Site	604.65 (1.70)	0.506 (2.21)
<i>JFQA</i> Web Site	825.16 (2.62)	-0.219 (-1.04)
<i>N</i>	213	191
Adjusted <i>R</i> <sup>2</sup>	97.67%	64.57%
Calendar year dummies	Yes	Yes
Journal-specific dummies	Yes	Yes

<sup>a</sup>n.a.: not available.

Subscription data are obtained for several leading finance and economics journals. Panel A of Table IV shows the data availability. For the *American Economic Review* (*AER*) and the *Journal of Finance*, subscription data are obtained from the Secretary’s Report in the Proceedings issue. For other journals, subscriptions are obtained from the Statement of Ownership, Management, and Circulation printed in that particular journal. The number of

subscriptions is defined to be the average number of total paid and/or requested circulation copies in the preceding 12 months. When subscription data are missing, I use the average of the two surrounding years.

To determine whether making forthcoming articles available affects subscriptions, I run the regression in Panel B of Table IV where the dependent variable is the number of subscribers. Both the *Journal of Finance* and the *Journal of Financial and Quantitative Analysis (JFQA)* made forthcoming articles available to the public in 1996. Although the *Journal of Finance* left the majority of papers online after publication, the *JFQA* removed articles once they appeared in print. I examine these practices by creating the variable *JF Web Site (JFQA Web Site)*, which equals one for the *Journal of Finance (JFQA)* after 1995 (1995), and zero otherwise. I include a variable that measures whether a journal is affiliated with an association, because more people may subscribe to journals whose subscription fee is part of the annual dues of an association. I also examine the real cost per page of the journal. This variable attempts to control for the journals' prices by dividing the annual cost (in 1994 dollars) by the number of pages that the journal published during the year. I control for general trends in subscriptions by including yearly dummy variables. Finally, the regression includes journal-specific effects to control for different subscription levels.

The *Web Site* coefficients are both positive, and the *JFQA* coefficient is significant. Thus, it appears that, at the very least, making forthcoming articles available to the public does not cause a decrease in subscriptions. Perhaps surprisingly, neither *Association Affiliation* nor *Cost Per Page* enter the regression significantly. The results are unchanged if I confine the analysis to the prior 20, or 10 years of data, or if I use a dummy variable for years after 1995 rather than annual dummies.

Panel B also includes a regression where the dependent variable is the SSCI impact factor, a measure of journal influence. In the impact factor regression, the *Web Site* dummy variables are set to one for the *Journal of Finance* and the *JFQA* after 1996. The one-year difference accounts for the fact that it takes one year for published papers to affect the SSCI impact factor.

The regression shows that the Web site does not appear to have any effect on the *JFQA* impact factor. In fact, the coefficient is negative. However, the coefficient on *JF Web Site* is positive and significant. The regression implies that posting forthcoming papers on the *Journal of Finance* Web site has increased the impact factor by 0.51. From 1977 to 1999, the average *Journal of Finance* impact factor is 1.58, indicating that the Web site has helped boost the impact factor by nearly one-third. The regression includes year and journal dummy variables; thus, the increase is not driven by an increasing trend or the fact that the *Journal of Finance* may have a larger impact factor than other journals. Rather, it appears that the significant increase may be attributable, at least in part, to the availability of articles via the Web site.

However, the coefficients must be interpreted cautiously. It may be too soon for inferences because impact factors measure cites to papers published in the prior two years. Because Web site papers were not published until

1996, only the last three years of impact factors measure the benefit of the Web site. This question should be revisited in a few years when the data is more complete.

## V. Conclusions

This paper analyzes the success of the *Journal of Finance* Web site at disseminating research. From November 1997 to November 1999, more than 284,000 papers were downloaded from the Web site. The average paper received nearly 85 downloads per month and was made available 311 days before appearing in print. Papers about “General Financial Markets” are downloaded more frequently than papers in other subject areas, while articles are downloaded more than shorter papers.

The lead article in the *Journal of Finance* is downloaded significantly more frequently than other papers for two reasons. First, the paper appears to be of greater interest because it is downloaded more often even before it is announced as the lead article. In addition, there appears to be a certification effect. Once a paper is announced as the lead, the marginal increase in downloads is significant. The results are weaker when looking at the lead shorter paper.

Downloads are positively correlated with future citations, supporting the use of downloads as another measure of research impact. On average, it takes 109 downloads per month to generate one citation per year. Finally, it does not appear that making forthcoming articles available to the public has an adverse effect on subscriptions. At the same time, availability may increase journal impact factors, but the true effect is only starting to surface.

In sum, the *Journal of Finance* Web site appears to disseminate finance research to a broad audience in an expedient manner. Additionally, the analysis gives greater insight into the responsibilities and power that editors have with regard to selecting the lead article in each issue.

Future research should examine downloads from other Web sites. Perhaps of greatest interest would be analyzing if the number of downloads from the Social Science Research Network (SSRN) is a good predictor of whether a paper becomes published and where. It seems clear that the use of downloads as a measure of impact is a new area which deserves further study.

## REFERENCES

- Borokhovich, Kenneth A., Robert J. Bricker, and Betty J. Simkins, 2000, An analysis of finance journal impact factors, *Journal of Finance* 55, 1457–1469.
- Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.
- Schwert, G. William, 1993, *The Journal of Financial Economics: A retrospective evaluation (1974 to 1991)*, *Journal of Financial Economics* 33, 369–424.
- Stulz, René M., 2000, Report of the editor of the *Journal of Finance* for the year 1999, *Journal of Finance* 55, 1861–1892.

