

Online Appendix for “*Investor Awareness or Information Asymmetry? Wikipedia and IPO Underpricing*”

ADDITIONAL ROBUSTNESS CHECKS

IPO offer price precision

Bradley, Cooney, Jordan, and Singh (2004) use integer offer prices as an indicator of low price precision and therefore as a measure of information asymmetry. To examine if a Wikipedia article helps to mitigate information asymmetry and improve price precision, we construct an indicator variable, *integer*, for IPOs priced on an integer and regress it on *Wikipedia*. Table A1, Model 1 reports that *Wikipedia* is not a significant determinant of integer offer prices. This provides additional evidence that the primary effect of a Wikipedia article is not a reduction in information asymmetry during the bookbuilding / price setting process.

Qualitative vs. quantitative information

Managers have incentives to delay the disclosure of bad news (Kothari, Shu, & Wysocki, 2009) and to publicize good news (Solomon, 2012). One technique managers apply to hide bad performance is to use “soft talk”, which is equivocal, unverifiable, and biased upward (Dambra, Wasley, & Wu, 2013). Consequently, markets react more strongly to numbers (Hutton, Miller, & Skinner, 2003). To examine the effect of quantitative information in Wikipedia articles and S-1 filings, we count numbers that start with a space and can consist of numeric characters (0-9), plus and minus signs (+ and -), currency symbols (\$€£), commas (,) and periods (.) and standardize the frequency of numbers by the total count of numbers and words in each document.

Results in Table A1, Model 2 show that a 1% increase in numbers in a company’s S-1 is associated with a 4.91 percentage point decrease of underpricing, consistent with the argument that more quantitative information can mitigate information asymmetry. The percentage of numbers in Wikipedia articles in Model 3, however, is not significant, suggesting that investor

attention from Wikipedia, rather than information asymmetry, is more related to IPO underpricing.

Underwriter promotion

Given that underwriters assume the risk of selling IPO shares, they are incentivized to act proactively by publicizing the IPO so as to avoid costly price stabilization in the secondary market (Ruud, 1993). If underwriters initiate Wikipedia articles for IPO companies to facilitate selling of shares, the existence of a pre-IPO Wikipedia article would not be exogenous.

According to the editing policy of Wikipedia, insider editing is discouraged although not completely prohibited. Wikipedia requests editors to reveal conflicts of interest that they have with the company and to first discuss edits on the article's "talk page" to get the approval of the Wikipedia community. Despite this request, conflict of interest editing has happened several times in Wikipedia's history. However, the rigorous detection mechanisms discussed previously help to ensure that biased content is removed promptly.¹

We also examine the time of article creation relative to the S-1 filing date. Of the 330 IPOs in the Wikipedia sample, 316 have a Wikipedia article prior to registration with the U.S. SEC. Figure A2 reports the distribution of differences between the S-1 filing date and date of Wikipedia article creation. We find that very few Wikipedia articles are created around the S-1 filing date, which alleviates the "self-promotion" concern.²

Because Wikipedia does not provide editor affiliation, we cannot directly observe whether a revision is made by an underwriter or other user. However, underwriter promotion is likely to be associated with the number of unique editors engaged in the production and revision of an IPO

¹ For example, Wikipedia investigated allegations that Bell Pottinger, the largest UK public relations firm, manipulated its clients' Wikipedia articles. See:

<https://www.independent.co.uk/news/uk/politics/wikipedia-founder-attacks-bell-pottinger-for-ethical-blindness-6273836.html>

² The results are qualitatively unchanged when we estimate our baseline results using the sample of firms with Wikipedia articles created before the IPO filing date.

firm's Wikipedia article. Therefore, we control for the unique number of editors and rerun Table 3, Model 3. These untabulated results show that the observed relationship between IPO underpricing and *Wikipedia* remains unchanged.

Long-run performance

We follow Ritter and Welch (2002) to examine long-run performance of IPOs. First, we compare the buy-and-hold abnormal return between the Wikipedia and non-Wikipedia samples. For each IPO, we calculate the buy-and-hold return (*BHR*) and buy-and-hold abnormal return (*BHAR*) from the first closing price to the earlier of the three-year IPO anniversary or the last available price on CRSP. Table A3, Panel A reports that the difference in *BHR* and *BHAR* between the Wikipedia and non-Wikipedia samples is not significant.

Due to the overlap of buy-and-hold returns, we conduct a time-series regression using the Fama and French (1993) three-factor model with the momentum factor (Carhart, 1997). The dependent variable is the equally-weighted monthly return (for 96 months from 2009 to 2016) in excess of the risk-free rate for the portfolio of IPOs that go public during the prior 36 months. We report the results in Table A3, Panel B. Neither the Wikipedia nor non-Wikipedia subsamples exhibit significant abnormal returns.³ However, non-Wikipedia IPOs have a relatively higher abnormal return compared with Wikipedia IPOs, consistent with the investor attention model prediction of Merton (1987) that “more widely-known firms with larger investor bases will have lower alphas.”

Corporate websites

Wikipedia is not the only online platform where an IPO company might receive investor attention. Company websites are one of the primary information sources for investors and we find that all sample firms have a corporate website prior to their IPO.⁴ Given the lack of cross-sectional

³ We also examine returns around IPO quiet period expiration in Figure A3 and Table A2.

⁴ The creation date of company website domain is obtained from <http://whois.domaintools.com>

variability, the existence of a company's website is not likely to explain the IPO underpricing difference between the Wikipedia and non-Wikipedia samples.

Emerging growth companies

We note that the JOBS Act was passed during our sample period with the goal of facilitating IPOs of emerging growth companies (EGCs). We examine whether Wikipedia has a different effect on underpricing for EGCs and non-EGCs. Because the JOBS Act was signed on April 5, 2012, we only include IPOs issued after that date. When we add an EGC indicator and its interaction with the *Wikipedia* indicator to our baseline regressions, we do not find a difference in the effect of Wikipedia on EGCs and non-EGCs (untabulated).

Wikipedia sentiment

We also consider the possibility that our results are driven by sentiment expressed in firms' Wikipedia articles. We examine differences between the language used in firms' S-1 registration statements and their Wikipedia articles. Given the different nature of Wikipedia articles and registration statements, we employ both the Loughran and McDonald (2011) (LM) dictionary, which is specifically designed for financial filings, and the Harvard General Inquirer (Harvard GI) dictionary, which is widely used in social science research. Loughran and McDonald (2013) find that underpricing is greater for firms with a high level of aggregate uncertainty in their S-1 filings.

In Table A5, Panel A we report the average percentage of words in each sentiment dictionary for both Wikipedia articles and S-1 filings for the 330 Wikipedia firms. Consistent with issuers' intention to lower litigation risk, S-1 filings are characterized by more negative tone (across both dictionaries), uncertainty, and litigious language use.⁵ Overall, these results indicate that Wikipedia and S-1 filings convey information differently.

⁵ We report the 20 words that appear most frequently and the corresponding percentage in each sentiment category for both the LM and Harvard GI dictionaries in Exhibit E.

To provide further insight on this relation, we regress underpricing on Wikipedia article sentiment. The results, which are reported in Table A5 Panel B, show that positive and net Harvard GI words are positively correlated with first-day returns. Thus, to the extent that Wikipedia articles with general positive tone increase investors' attention to the IPO, these results are consistent with the investor attention hypothesis. The lack of significance for the LM dictionary variables suggests that underpricing is more impacted by the general content within Wikipedia articles as opposed to financial context-specific content that is more prevalent in S-1 filings.⁶

⁶ We find that sentiment differences between S-1 filings and Wikipedia articles have no relation to underpricing. Table A1 shows no (negative) relationship between IPO underpricing and *Wikipedia* (S-1) quantitative information.

EXHIBIT A – WIKIPEDIA ARTICLE IDENTIFICATION DETAILS

We use a web crawler to search for an IPO firm’s Wikipedia article and manually check its accuracy. We assign a Wikipedia article to an IPO if the article is titled with the name of:

(1) the IPO firm;

(2) the IPO firm’s parent company if it is the IPO firm’s parent before the first trading date;

For example, Bruker AXS Inc. is an IPO firm in our sample. The Wikipedia article we record is that of its parent. (<https://en.wikipedia.org/wiki/Bruker>). Other examples include Carolina Group-Loews Corporation and Travelers Property Casualty-Citigroup. For Delek US Holdings Inc.’s page we use “Delek”

(3) the IPO firm’s major subsidiary;

For example, for Hertz Global Holdings Inc. we use the Wikipedia article of The Hertz Corporation. For NYMEX Holdings Inc., we use the Wikipedia article of New York Mercantile Exchange

(4) a company from which the IPO firm separates;

For example, Reliant Energy separated into CenterPoint Energy, Inc. and Reliant Resources, Inc., which is an IPO firm in our sample. The URL of the Wikipedia article is https://en.wikipedia.org/wiki/GenOn_Energy.

(5) the firm’s predecessor;

For example, for Bakers Footwear Group Inc. we use the Wikipedia article of Edison Brothers Stores, Inc.


(6) the core product or service and primarily contains information about the firm.

For example, for Intersections Inc., we use the Wikipedia article of its service “Identity guard.” Neurometrix Inc. uses the article “Quell” ([https://en.wikipedia.org/wiki/Quell_\(wearable\)](https://en.wikipedia.org/wiki/Quell_(wearable))). Lincoln Educational Services Corp has an article titled “Lincoln Group of Schools”

Some firms do not have an exclusive article but have brief information on a page with other items that are classified as “ambiguous words” by Wikipedia. For example, Veridian Corp is on a page titled “Veridian.” On this page, Veridian Corp. is the first item and the contents include “an American aerospace and defense company, acquired by General Dynamics in 2003. Veridian Engineering, Inc., a subsidiary of American aerospace and defense company Veridian Corporation which was acquired by General Dynamics in 2003.” <https://en.wikipedia.org/wiki/Veridian>. For these instances, we do not consider the firm to have a Wikipedia article because much of our focus is on the information provided by Wikipedia instead of simply the mention of the firm in a page.

For firms with a Wikipedia article prior to the first-trading day lacking content or with less than 30 words in the main body, we set the Wikipedia indicator equal to zero. For example, Allot Communications Ltd went public on Nov 15, 2006 and its Wikipedia article was created on Dec 9, 2005. However, the article was a redirect page to “Allotment”. The first revision after its creation was on Sept 17, 2007. Only one IPO firm has its Wikipedia article created on its IPO date (Zoetis Inc.).

EXHIBIT B – “ARTICLE INFORMATION” PAGE OF LINKEDIN



WIKIPEDIA
The Free Encyclopedia


- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

- What links here
- Related changes
- Upload file
- Special pages
- Page information
- Wikidata item


Languages 

Article [Talk](#)

Information for "LinkedIn"

[Help:Page information](#)

Basic information

Display title	LinkedIn
Default sort key	LinkedIn
Page length (in bytes)	84,426
Page ID	970755
Page content language	en - English
Page content model	wikitext
Indexing by robots	Allowed
Number of page watchers	437
Number of page watchers who visited recent edits	49
Number of redirects to this page	25
Counted as a content page	Yes
Wikidata item ID	Q213660
Central description	social networking website for people in professional occupations
Page image	
Page views in the past 30 days	95,791

Page protection

Edit	Allow all users (no expiry set)
Move	Allow all users (no expiry set)

Edit history

Page creator	AAAAA (talk contribs)
Date of page creation	04:09, 9 September 2004
Latest editor	Titusmars (talk contribs)
Date of latest edit	11:13, 21 August 2018
Total number of edits	2,638
Recent number of edits (within past 30 days)	19
Recent number of distinct authors	11

via <https://en.wikipedia.org/w/index.php?title=LinkedIn&action=info>

EXHIBIT C – WIKIPEDIA ARTICLE OF LINKEDIN ON MAY 19TH, 2011



- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

- Interaction
- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

- Tools
- What links here
- Related changes
- Upload file
- Special pages
- Permanent link
- Page information
- Wikidata item
- Cite this page
- Print/export
- Create a book
- Download as PDF
- Printable version

Article [Talk](#)

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)


[Read](#) [Edit](#) [View history](#)

Search Wikipedia

LinkedIn

From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by **Wikidemon** ([talk](#) | [contribs](#)) at 07:55, 11 May 2012 (*Undo revision 491943240 by 31.3.241.40 (talk)*). The present address (URL) is a **permanent link** to this revision, which may differ significantly from the **current revision**. ([diff](#)) [←](#) [Previous revision](#) | [Latest revision \(diff\)](#) | [Newer revision](#) [→](#) ([diff](#))

 This article may **require cleanup** to meet Wikipedia's quality standards. No cleanup reason has been specified. Please help improve this article if you can. *(August 2011)* ([Learn how and when to remove this template message](#))

LinkedIn (NYSE: LNKD) (/[invalid input: 'icon'], lɪŋkɪn.ɪn) is a business-related *social networking* site. Founded in December 2002 and launched in May 2003,^[4] it is mainly used for *professional networking*. As of 9 February 2012, LinkedIn reports more than 150 million registered users in more than 200 countries and territories.^{[5][6]} The site is available in English, French, German, Italian, Portuguese, Spanish, Swedish, Romanian, Russian, Turkish, Japanese, Czech and Polish.^{[7][8][9]} Quantcast reports LinkedIn has 21.4 million monthly unique U.S. visitors and 47.6 million globally.^[10] In June 2011, LinkedIn had 33.9 million unique visitors, up 63 percent from a year earlier and surpassing MySpace.^[11] LinkedIn filed for an *initial public offering* in January 2011 and traded its first shares on May 19, 2011, under the NYSE symbol "LNKD".^[12]

Contents [hide]

- History
- Membership
- Features
 - Applications
 - Mobile
 - Groups
 - Job listings
- Reception
- International reception
- SNA LinkedIn
- See also
- References
- External links

LinkedIn Corporation

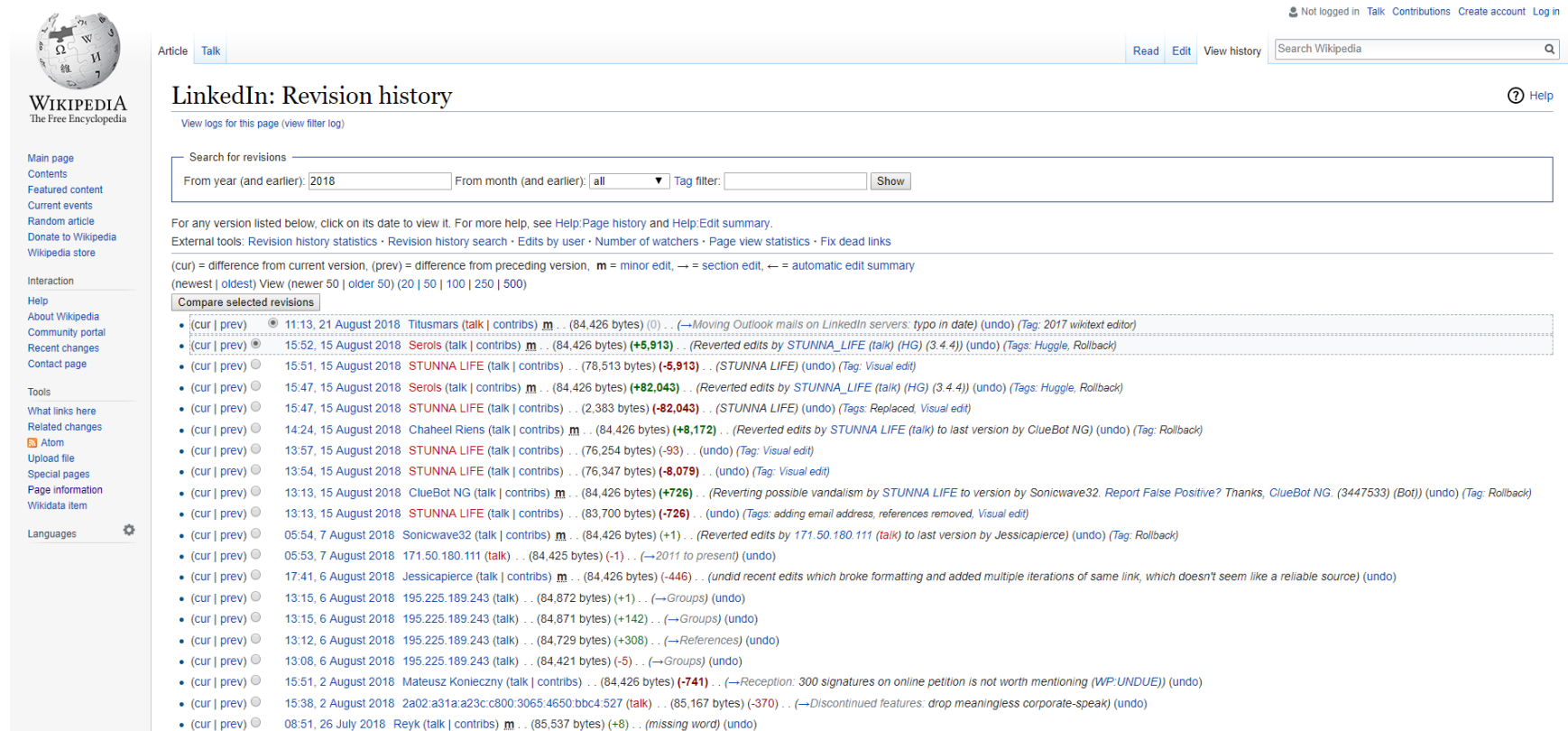


File:LinkedInHomepage.PNG
LinkedIn homepage as of July 2011

Type of business	Public
Type of site	Social network service
Available in	English, French, German, Dutch, Italian, Portuguese, Spanish, Swedish, Romanian, Russian, Turkish, Japanese, Czech and Polish.
Traded as	NYSE: LNKD
Founded	Santa Monica, California (2003)
Headquarters	Mountain View, California, US
Area served	Worldwide
Founder(s)	Reid Hoffman Allen Blue

via <https://en.wikipedia.org/w/index.php?title=LinkedIn&oldid=491963773>

EXHIBIT D.1 – REVISION HISTORY PAGE OF LINKEDIN ARTICLE




The screenshot shows the Wikipedia revision history page for the article "LinkedIn". The page title is "LinkedIn: Revision history". The current version is 11:13, 21 August 2018, by Titusmars. The page size is 84,426 bytes. The revision history is a list of 25 revisions, each with a date, time, and user name. The most recent revision is 11:13, 21 August 2018, by Titusmars, which is the current version. The second most recent revision is 15:52, 15 August 2018, by Serols, which is a minor edit (+5,913 bytes) that reverted edits by STUNNA_LIFE. The third most recent revision is 15:51, 15 August 2018, by STUNNA_LIFE, which is a visual edit (-5,913 bytes). The fourth most recent revision is 15:47, 15 August 2018, by Serols, which is a minor edit (+82,043 bytes) that reverted edits by STUNNA_LIFE. The fifth most recent revision is 15:47, 15 August 2018, by STUNNA_LIFE, which is a replaced edit (-82,043 bytes). The sixth most recent revision is 14:24, 15 August 2018, by Chaheel Riens, which is a minor edit (+8,172 bytes) that reverted edits by STUNNA_LIFE to the last version by ClueBot NG. The seventh most recent revision is 13:57, 15 August 2018, by STUNNA_LIFE, which is a visual edit (-93 bytes). The eighth most recent revision is 13:54, 15 August 2018, by STUNNA_LIFE, which is a visual edit (-8,079 bytes). The ninth most recent revision is 13:13, 15 August 2018, by ClueBot NG, which is a minor edit (+726 bytes) that reverted possible vandalism by STUNNA_LIFE to the version by Sonicwave32. The tenth most recent revision is 13:13, 15 August 2018, by STUNNA_LIFE, which is a minor edit (-726 bytes) that added an email address and removed references. The eleventh most recent revision is 05:54, 7 August 2018, by Sonicwave32, which is a minor edit (+1 byte) that reverted edits by 171.50.180.111 to the last version by Jessicapierce. The twelfth most recent revision is 05:53, 7 August 2018, by 171.50.180.111, which is a minor edit (-1 byte) that moved the page to the present. The thirteenth most recent revision is 17:41, 6 August 2018, by Jessicapierce, which is a minor edit (-446 bytes) that undid recent edits which broke formatting and added multiple iterations of same link, which doesn't seem like a reliable source. The fourteenth most recent revision is 13:15, 6 August 2018, by 195.225.189.243, which is a minor edit (+1 byte) that added a group. The fifteenth most recent revision is 13:15, 6 August 2018, by 195.225.189.243, which is a minor edit (+142 bytes) that added a group. The sixteenth most recent revision is 13:12, 6 August 2018, by 195.225.189.243, which is a minor edit (+308 bytes) that added references. The seventeenth most recent revision is 13:08, 6 August 2018, by 195.225.189.243, which is a minor edit (-5 bytes) that added a group. The eighteenth most recent revision is 15:51, 2 August 2018, by Mateusz Konieczny, which is a minor edit (-741 bytes) that received 300 signatures on an online petition. The nineteenth most recent revision is 15:38, 2 August 2018, by 2a02:a31a:a23c:c800:3065:4650:bbc4:527, which is a minor edit (-370 bytes) that discontinued features. The twentieth most recent revision is 08:51, 26 July 2018, by Reyk, which is a minor edit (+8 bytes) that added a missing word.

via <https://en.wikipedia.org/w/index.php?title=LinkedIn&action=history>

A full list of edits to LinkedIn's Wikipedia article during its IPO in May 2011 are described in detail at:

<https://en.wikipedia.org/w/index.php?title=LinkedIn&action=history&year=2011&month=5>

EXHIBIT D.2 – COMPARISON BETWEEN HISTORICAL WIKIPEDIA ARTICLES OF LINKEDIN



Not logged in | [Talk](#) | [Contributions](#) | [Create account](#) | [Log in](#)

Article Talk
Read Edit View history

LinkedIn: Difference between revisions

From Wikipedia, the free encyclopedia

Browse history interactively

Revision as of 13:15, 6 August 2018 (edit)

195,225,189,243 (talk)

[\(←Groups\)](#)
[← Previous edit](#)

Latest revision as of 11:13, 21 August 2018 (edit) (undo)

Titusmars (talk | contribs)

m (←Moving Outlook mails on LinkedIn servers: typo in date)
(Tag: 2017 wikitext editor)

(13 intermediate revisions by 8 users not shown)

Line 328:

=== Groups ===

LinkedIn also supports the formation of interest groups, and as of March 29, 2012 there are 1,248,019 such groups whose membership varies from 1 to 744,662.<ref name=autogenerated1>{{cite web|url=http://www.linkedin.com/groupsDirectory?results=&ppiSearchOrigin=GLHD&keywords=|title=Groups Directory |publisher=LinkedIn |accessdate=December 8, 2011}}</ref><ref>{{cite news|publisher=i-newswire.com|url=http://www.i-newswire.com/world-s-largest-linkedin-group/160028 |title= World's Largest LinkedIn Group Breaks The 700,000 Member mark |accessdate=February 28, 2012}}</ref> The majority of the largest groups are employment related, although a very wide range of topics are covered mainly around professional and career issues, and, at 6th of August 2018, there are over 88,000 groups for both academic and corporate alumni.<ref>{{cite web|url=https://www.immigrantschool.co.uk/amp/linkedin-and-your-career|title=LinkedIn and your career |publisher=Immigrant School |accessdate=August 6,2018}}</ref> Groups support a limited form of discussion area, moderated by the group owners and managers.<ref>{{cite web|url=http://socialmediatoday.com/mark-lerner/2070731/how-avoid-linkedin-s-site-wide-automatic-moderation|title=How to Avoid LinkedIn's Site Wide Automatic Moderation (SWAM)|website=Socialmediatoday.com|accessdate=October 17, 2014|archive-uri=https://web.archive.org/web/20140627130825/http://socialmediatoday.com/mark-lerner/2070731/how-avoid-linkedin-s-site-wide-automatic-moderation|archive-date=June 27, 2014|dead-uri=yes|df=mdy-all}}</ref> Since groups offer the functionality to reach a wide audience without so easily falling foul of [[Anti-spam techniques|anti-spam solutions]], there is a constant stream of spam postings, and there now exist a range of firms who offer a spamming service for this very purpose. LinkedIn has devised a few mechanisms to reduce the volume of spam.<ref>{{cite web|last=Lerner|first=Mark|title=How To Avoid LinkedIn's Site Wide Automatic Moderation|url=http://blog.oktopost.com/how-to-avoid-linkedin-swam-site-wide-automatic-moderation|publisher=Oktopost|accessdate=24 March 2014|archive-uri=https://web.archive.org/web/20140324171633/http://blog.oktopost.com/how-to-avoid-linkedin-swam-site-wide-automatic-moderation|archive-date=March 24, 2014|dead-uri=yes|df=mdy-all}}</ref> but recently{{When|date=April 2016}} took the decision to remove the ability of group owners to inspect the email address of new members in order to determine if they were spammers.{{Citation needed|date=April 2016}} Groups also keep their members informed through emails with updates to the group, including most talked about

Line 328:

=== Groups ===

LinkedIn also supports the formation of interest groups, and as of March 29, 2012 there are 1,248,019 such groups whose membership varies from 1 to 744,662.<ref name=autogenerated1>{{cite web|url=http://www.linkedin.com/groupsDirectory?results=&ppiSearchOrigin=GLHD&keywords=|title=Groups Directory |publisher=LinkedIn |accessdate=December 8, 2011}}</ref><ref>{{cite news|publisher=i-newswire.com|url=http://www.i-newswire.com/world-s-largest-linkedin-group/160028 |title= World's Largest LinkedIn Group Breaks The 700,000 Member mark |accessdate=February 28, 2012}}</ref> The majority of the largest groups are employment related, although a very wide range of topics are covered mainly around professional and career issues, and there are **currently**{{When|date=April 2016}} 128,000 groups for both academic and corporate alumni.{{Citation needed|date=April 2016}} Groups support a limited form of discussion area, moderated by the group owners and managers.<ref>{{cite web|url=http://socialmediatoday.com/mark-lerner/2070731/how-avoid-linkedin-s-site-wide-automatic-moderation|title=How to Avoid LinkedIn's Site Wide Automatic Moderation (SWAM)|website=Socialmediatoday.com|accessdate=October 17, 2014|archive-uri=https://web.archive.org/web/20140627130825/http://socialmediatoday.com/mark-lerner/2070731/how-avoid-linkedin-s-site-wide-automatic-moderation|archive-date=June 27, 2014|dead-uri=yes|df=mdy-all}}</ref> Since groups offer the functionality to reach a wide audience without so easily falling foul of [[Anti-spam techniques|anti-spam solutions]], there is a constant stream of spam postings, and there now exist a range of firms who offer a spamming service for this very purpose. LinkedIn has devised a few mechanisms to reduce the volume of spam.<ref>{{cite web|last=Lerner|first=Mark|title=How To Avoid LinkedIn's Site Wide Automatic Moderation|url=http://blog.oktopost.com/how-to-avoid-linkedin-swam-site-wide-automatic-moderation|publisher=Oktopost|accessdate=24 March 2014|archive-uri=https://web.archive.org/web/20140324171633/http://blog.oktopost.com/how-to-avoid-linkedin-swam-site-wide-automatic-moderation|archive-date=March 24, 2014|dead-uri=yes|df=mdy-all}}</ref> but recently{{When|date=April 2016}} took the decision to remove the ability of group owners to inspect the email address of new members in order to determine if they were spammers.{{Citation needed|date=April 2016}} Groups also keep their members informed through emails with updates to the group, including most talked about discussions within your professional circles.<ref name=autogenerated1 /></ref>{{cite

via <https://en.wikipedia.org/w/index.php?title=LinkedIn&type=revision&diff=855870683&oldid=853702839>

EXHIBIT E - TOP WORDS IN WIKIPEDIA ARTICLES AND S-1 FILINGS

Panel A. Wikipedia articles

<u>positive</u>		<u>negative</u>		<u>uncertainty</u>		<u>litigious</u>		<u>positive GI</u>		<u>negative GI</u>	
word	%	word	%	word	%	word	%	word	%	word	%
BEST	10.4	ERRORS	2.8	APPROXIMATELY	10.0	CONTRACT	8.9	HEALTH	4.9	DEAD	5.5
GREAT	3.3	AGAINST	2.6	MAY	8.8	CLAIMS	6.6	HOME	4.7	DIVISION	4.8
LEADING	3.2	LATE	2.1	RISK	8.8	LAWSUIT	6.5	BEST	4.4	AGAINST	3.6
BETTER	2.9	CLOSED	2.0	COULD	8.6	COURT	5.6	ENTERTAINMENT	2.3	YELP	2.9
POPULAR	2.7	CLAIMS	1.8	UNKNOWN	4.9	LEGAL	5.2	PROTECTION	1.8	BANKRUPTCY	2.5
WINNER	2.6	BANKRUPTCY	1.8	POSSIBLE	4.9	LAW	5.0	SUPER	1.8	CANCER	1.9
LEADERSHIP	2.4	SPAM	1.7	NEARLY	4.4	CONTRACTS	4.5	COMMUNITY	1.7	INVALID	1.9
WINNERS	2.3	LACKING	1.6	SUGGESTED	3.3	SETTLEMENT	3.5	EDUCATION	1.6	LOSS	1.8
INNOVATION	2.2	FORCE	1.3	ALMOST	2.9	DOCKET	3.3	GRAND	1.6	CONTROVERSY	1.6
ABLE	2.1	INVALID	1.3	MIGHT	1.7	CLAIM	2.3	PARTNERSHIP	1.4	HUNGRY	1.6
GOOD	2.1	LOSS	1.3	DEPENDING	1.7	REGULATORY	2.3	FRESH	1.4	THEFT	1.4
SUCCESSFUL	2.0	IGNORED	1.2	ROUGHLY	1.6	SUED	2.2	PRIVACY	1.4	ERROR	1.4
ALLIANCE	1.9	CONTROVERSY	1.2	SOMETIMES	1.3	BREACH	1.9	CREATE	1.3	EMERGENCY	1.4
SUCCESS	1.8	ERROR	1.0	VARIABLE	1.2	LAWS	1.5	INTELLIGENCE	1.3	COMPETITOR	1.4
ENABLES	1.4	CUT	0.9	VARY	1.1	LAWSUITS	1.4	PARTNER	1.2	COMPETITION	1.4
INNOVATIVE	1.3	LOST	0.8	BELIEVED	1.0	AMENDMENT	1.4	SOLUTION	1.2	SAP	1.1
EXCLUSIVE	1.3	FRAUD	0.8	APPEARS	0.9	JUSTICE	1.3	POPULAR	1.1	WAR	1.1
GREATER	1.3	PROBLEMS	0.7	SPECULATION	0.9	ALLEGED	1.3	HUMAN	1.0	DEATH	1.1
POSITIVE	1.2	CRITICISM	0.7	AMBIGUOUS	0.9	REGULATORS	1.3	ABILITY	0.9	COMPETE	1.1
EASY	1.1	CONCERNS	0.7	APPEARED	0.9	LITIGATION	1.0	FITNESS	0.9	FRAUD	1.1

Panel B. S-1 filings

<u>positive</u>		<u>negative</u>		<u>uncertainty</u>		<u>litigious</u>		<u>positive GI</u>		<u>negative GI</u>	
word	%	word	%	word	%	word	%	word	%	word	%
EFFECTIVE	10.2	LOSS	5.3	MAY	33.1	AMENDED	6.1	SIGNIFICANT	5.6	LOSS	8.9
BENEFIT	5.2	AGAINST	2.9	COULD	12.6	REGULATORY	5.9	EFFECTIVE	4.7	LIABILITY	6.2
ABLE	5.1	CLAIMS	2.8	APPROXIMATELY	7.5	LAWS	5.8	ABILITY	4.4	AGAINST	4.8
GREATER	3.4	ADVERSELY	2.7	BELIEVE	6.3	REGULATIONS	4.9	PRO	3.1	ADVERSE	3.8
GAIN	2.6	RESTATED	2.6	RISK	4.6	CLAIMS	4.5	ABLE	2.3	COMPETITIVE	2.9
BENEFICIAL	2.4	LOSSES	2.6	RISKS	2.8	LAW	4.3	OBTAIN	2.3	FAILURE	2.6
SUCCESSFUL	2.2	ADVERSE	2.3	ASSUMPTIONS	1.9	CONTRACT	3.7	HEALTH	2.1	EXCESS	2.5
SUCCESS	2.1	TERMINATION	2.2	ASSUMED	1.7	CONTRACTS	3.5	RELEVANT	1.6	DEPRECIATION	2.1
OPPORTUNITIES	2.1	CLOSING	2.0	INTANGIBLE	1.4	LEGAL	3.4	PROPRIETARY	1.4	DIFFICULT	1.8
ACHIEVE	2.0	IMPAIRMENT	1.8	ASSUMING	1.2	SHALL	3.0	BONUS	1.3	COMPETITION	1.8
SUCCESSFULLY	2.0	FAILURE	1.6	MIGHT	1.1	INDEMNIFICATION	2.7	PARTNER	1.2	LIQUIDATION	1.8
GOOD	1.8	UNABLE	1.6	ANTICIPATE	1.1	AMENDMENT	2.3	REASONABLE	1.1	COMPETE	1.7
BENEFICIALLY	1.7	LITIGATION	1.4	POSSIBLE	1.1	LITIGATION	2.2	SUFFICIENT	1.1	CANCER	1.4
OPPORTUNITY	1.7	LIMITATIONS	1.3	DEPEND	1.0	CONSENT	1.8	PARTNERSHIP	1.1	DISEASE	1.4
PROFITABILITY	1.7	TERMINATE	1.3	ANTICIPATED	0.9	CONTRACTUAL	1.7	OFFSET	1.1	VOLATILITY	1.3
LEADING	1.7	DECLINE	1.2	VOLATILITY	0.9	REGULATION	1.6	BENEFICIAL	1.1	BREACH	1.3
BEST	1.7	DIFFICULT	1.1	PENDING	0.9	SETTLEMENT	1.5	COMPREHENSIVE	1.1	DIFFER	1.3
ENABLE	1.5	DELAY	1.1	DIFFER	0.8	CLAIM	1.5	AUTHORITY	1.0	DEFICIT	1.1
FAVORABLE	1.5	TERMINATED	1.1	FLUCTUATIONS	0.8	COURT	1.4	PROTECTION	1.0	NEGATIVE	1.1
IMPROVE	1.4	LIQUIDATION	1.1	VARIABLE	0.7	BREACH	1.2	SUCCESSFUL	1.0	LIMITATION	1.0

FIGURE A1 – INDUSTRY DISTRIBUTION OF IPO FIRMS

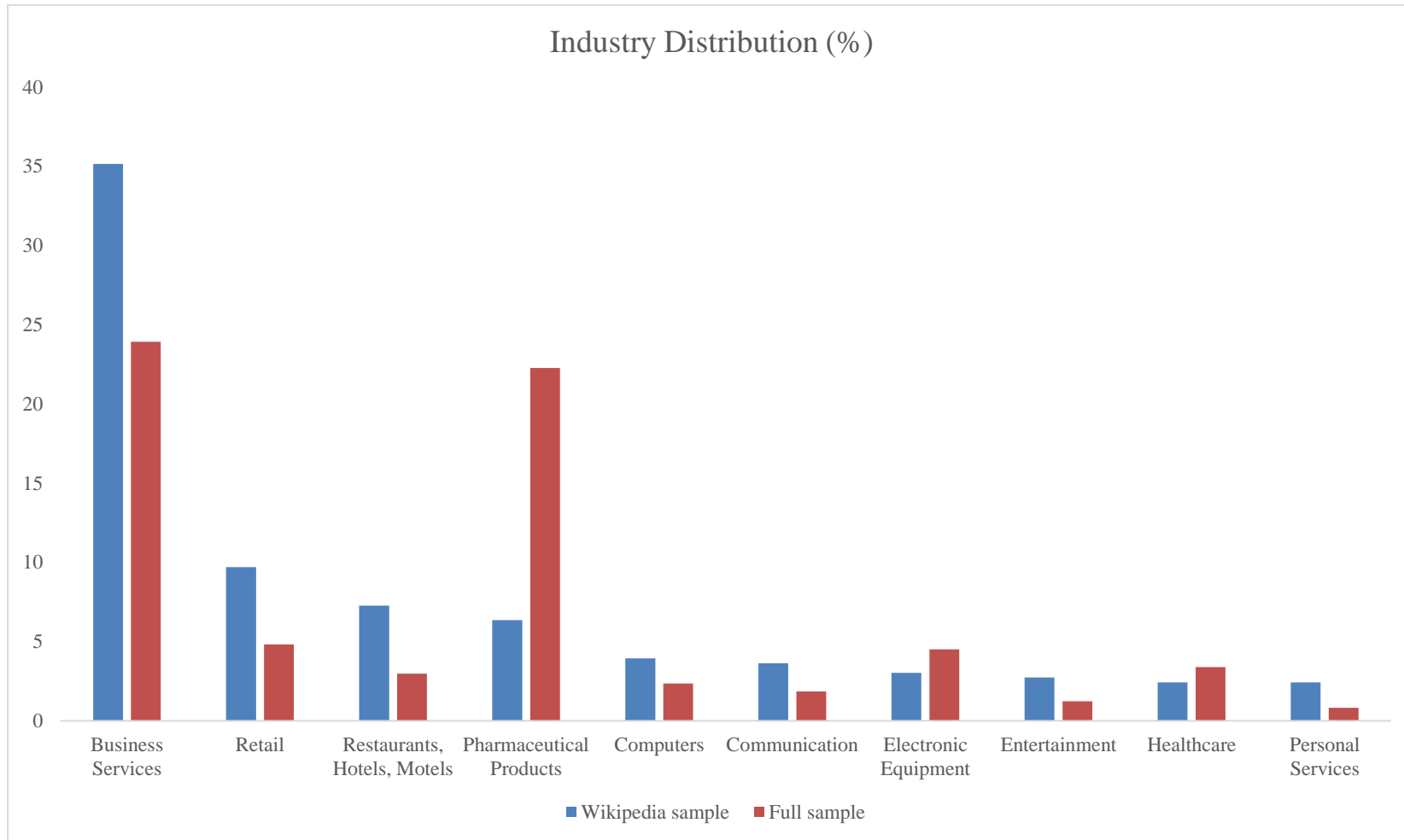


Figure A1 shows the Top 10 most represented industries in the sample of IPO firms with a pre-IPO Wikipedia article and the corresponding proportion of each industry in the full sample (i.e., 974 IPOs). The vertical axis is the proportion of IPOs in each industry, displayed in percentage.

FIGURE A2 – TIME DIFFERENCES BETWEEN WIKIPEDIA ARTICLE CREATION AND S-1 FILING (IN DAYS)

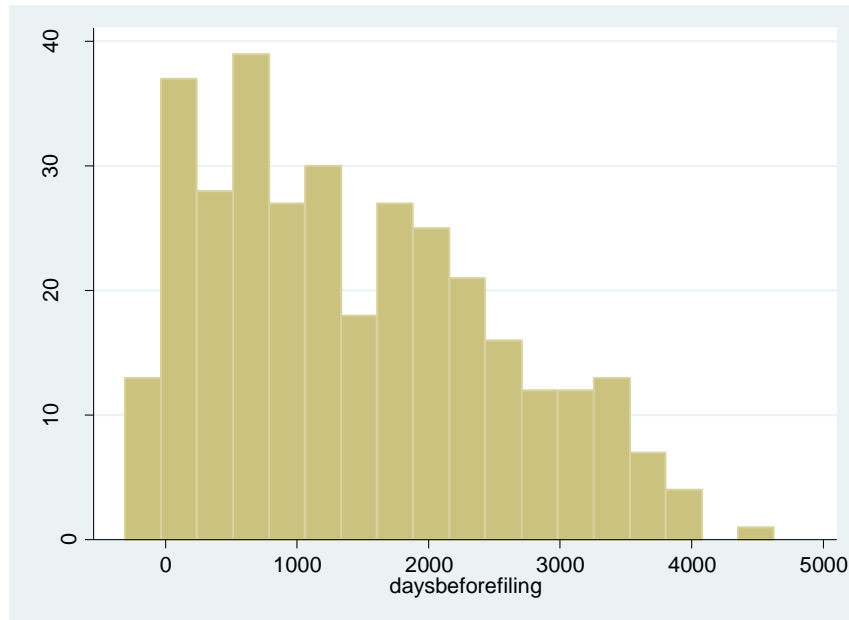


Figure A2 reports the distribution of the number of days between the S-1 filing date and the Wikipedia article creation date for IPO firms. The vertical axis is the proportion of IPOs in each time interval.

FIGURE A3 – CUMULATIVE ABNORMAL RETURNS: WIKIPEDIA AND NON-WIKIPEDIA SAMPLES

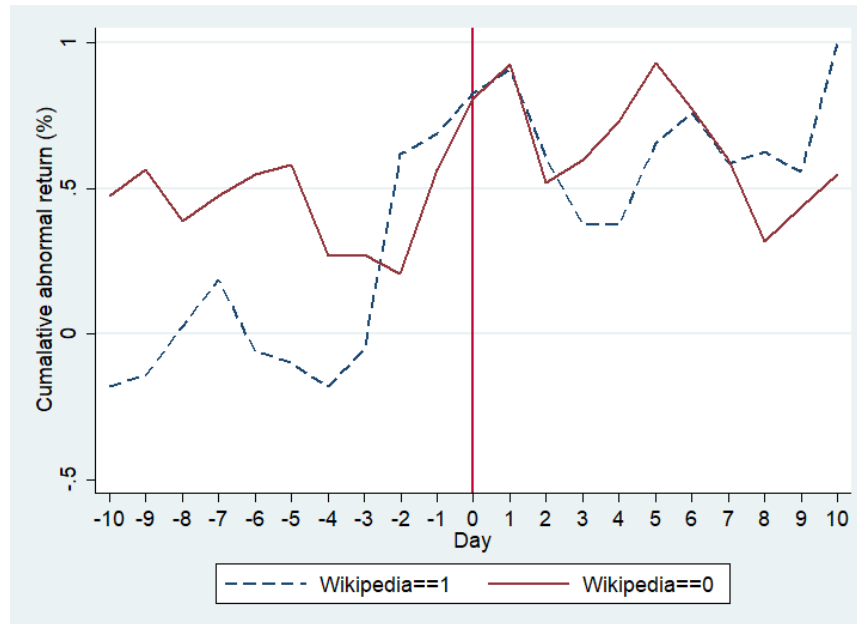


Figure A2 reports cumulative abnormal returns (CARs) in a [-10,+10] day window around the IPO quiet period expiration date. The horizontal axis measures the day relative to quiet period expiration date (Day 0). The vertical axis is CAR.

TABLE A1 – ADDITIONAL ROBUSTNESS CHECKS

Table A1 reports regression results for various robustness checks. *integer* is an indicator variable that is equal to 1 if the offer price is an integer and zero otherwise. *pctnum_s1* and *pctnum_wiki* are the percentage of numbers in the firm S-1 filing and Wikipedia article, respectively. The Appendix provides variable definitions. Standard errors in parentheses are robust and adjusted for clustering by Fama-French (1997) 48-industry and year. ***, **, * indicate significance at the 1%, 5%, and 10% level, respectively.

VARIABLES	(1) integer	(2) underpricing	(3) underpricing
Wikipedia	0.023 (0.111)	5.222* (2.344)	
pctnum_s1		-4.916*** (1.224)	
pctnum_wiki			-0.682 (0.965)
VC	0.419*** (0.109)	6.467** (2.633)	12.714*** (3.848)
top_tier	0.108 (0.188)	5.830*** (1.175)	10.697 (8.654)
overhang	-0.004 (0.021)	1.385*** (0.278)	3.021*** (0.464)
pos_EPS	-0.127 (0.098)	3.702 (2.494)	4.607 (4.374)
log_sales	0.023 (0.047)	0.043 (0.591)	-1.725 (0.987)
nasdaq15	-0.002 (0.024)	0.347 (0.300)	1.004* (0.494)
tech	-0.397*** (0.055)	2.565 (2.900)	-5.479 (4.202)
log_age	0.094 (0.106)	-0.998 (1.454)	2.132 (2.870)
log_news	-0.010 (0.059)	2.377** (0.800)	3.496** (1.305)
Constant	1.121*** (0.308)	18.260** (6.137)	-7.908 (15.061)
Observations	974	974	330
Year FE	Yes	Yes	Yes
Adjusted / Pseudo R^2	0.097	0.140	0.155

TABLE A2 – POST-IPO STOCK PERFORMANCE

Table A2 shows multivariate analysis of cumulative returns during the [+1,+12] month window after IPO. *cret* is [+1,+12] month cumulative IPO return and *cret_ind_adj* is [+1,+12] monthly cumulative IPO return adjusted by corresponding industry return. *d_sentiment* is the monthly investor sentiment change (orthogonal to macro variables) at the month the firm goes public; *cret_ind_pre* is the cumulative industry return over the [-6,-1] window. Unreported controls include log transformed offering proceeds, firm age, and assets. Standard errors (in parentheses) are clustered by offering year and quarter. ***, **, * indicate significance at the 1%, 5%, and 10% level, respectively.

VARIABLES	(1) cret	(2) cret	(3) cret	(4) cret	(5) cret	(6) cret
underpricing	0.038 (0.122)	0.083 (0.227)	0.138 (0.163)	0.119 (0.174)	0.016 (0.090)	0.318 (0.297)
Wikipedia	1.561 (3.183)	0.165 (3.784)	2.149 (3.414)	2.236 (3.303)	0.992 (2.918)	-1.745 (2.775)
Wikipedia × underpricing		0.082 (0.254)				-0.237 (0.261)
log_news	4.366 (2.858)	4.294 (2.931)	6.105 (3.925)	4.430 (2.875)	4.485 (2.836)	5.875 (4.072)
log_news × underpricing			-0.100 (0.197)			-0.090 (0.202)
offer_revision	-0.409 (0.228)	-0.408 (0.269)	-0.405 (0.229)	-0.318 (0.371)	-0.396 (0.270)	-0.307 (0.356)
offer_revision × underpricing				-0.007 (0.013)		-0.007 (0.014)
d_sentiment	-0.027 (0.794)	-0.034 (0.785)	-0.049 (0.807)	-0.052 (0.850)	-0.427 (0.851)	-0.407 (0.867)
d_sentiment × underpricing					0.023 (0.023)	0.019 (0.025)
top_tier	20.779 (9.781)	20.853 (9.734)	20.822 (9.808)	20.398 (9.908)	20.190 (9.558)	20.197 (9.753)
VC	-3.803 (6.168)	-3.679 (6.126)	-3.739 (6.193)	-3.735 (6.012)	-3.418 (6.288)	-3.016 (6.153)
overhang	-1.422 (1.047)	-1.470 (1.181)	-1.327 (1.238)	-1.384 (1.060)	-1.367 (1.045)	-1.394 (1.297)
cret_pre_ind	-0.472 (0.674)	-0.473 (0.631)	-0.468 (0.633)	-0.469 (0.595)	-0.466 (0.634)	-0.464 (0.612)
N	831	831	831	831	831	831
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.027	0.026	0.028	0.028	0.028	0.028

TABLE A3 – IPO LONG-RUN PERFORMANCE

Table A3 reports the results of IPO long-run performance. Panel A reports the three-year buy-and-hold return for IPOs from 2006 to 2013. *BHR* is the buy-and-hold return. *BHRM* is the compounded daily return of the CRSP value-weighted index. *BHAR* is the difference between *BHR* of IPOs and compounded daily return of the CRSP value-weighted index. Panel B reports calendar-time approach regression results of equally weighted monthly percentage returns on a portfolio of IPOs that have gone public during the prior 36 months from 2009 to 2016 for Wikipedia and non-Wikipedia samples, respectively. The sample period ends in 2013 to ensure that the long-run performance window is consistent for IPOs through all years. One issue with calendar-time regression is look-ahead bias. Specifically, an IPO that stops trading prior to its three-year anniversary is excluded from the sample. For example, Traffic.com went public on Jan 25, 2006 and the last available return is on March 6, 2007.

Panel A: Wikipedia cross-section average three-year IPO buy-and-hold return

	Wikipedia (N=214)	non-Wikipedia (N=427)	Difference	t-stat
BHR	24.95	10.77	14.18	1.50
BHAR	-3.20	-2.22	-0.98	-0.11

Panel B: Wikipedia cross-section multifactor regression of equally-weighted IPO portfolio returns

VARIABLES	Wikipedia			non-Wikipedia		
	(1) ret_rf	(2) ret_rf	(3) ret_rf	(4) ret_rf	(5) ret_rf	(6) ret_rf
Constant	-0.165 (0.409)	-0.167 (0.305)	-0.184 (0.292)	0.068 (0.520)	0.056 (0.413)	0.034 (0.398)
mktrf	1.310*** (0.094)	1.158*** (0.082)	1.110*** (0.080)	1.375*** (0.120)	1.208*** (0.111)	1.146*** (0.109)
smb		1.136*** (0.134)	1.144*** (0.128)		1.312*** (0.182)	1.323*** (0.175)
hml		-0.329*** (0.115)	-0.438*** (0.116)		-0.406** (0.156)	-0.548*** (0.158)
mom			-0.185*** (0.062)			-0.239*** (0.084)
N	96	96	96	96	96	96
R ²	0.673	0.825	0.841	0.584	0.747	0.768

TABLE A4 – QUIET PERIOD EXPIRATION EVENT STUDY RESULTS

Table A4 reports the results of market reaction at IPO quiet period expiration date. Panel A reports the average market-adjusted return (MAR) in each day of a [-10,+10] event window. Panel B reports the cumulative market-adjusted returns (CMAR) in 6 different event windows. The Wikipedia sample has a significant MAR on Day -2, while the non-Wikipedia sample has stronger fluctuations but a weak overall magnitude. Examining multi-day windows, we find that CMAR is significant for both samples over the [-2,+1] window; however, results are different depending on the event window. In sum, we do not observe a strong market reaction around the quiet period expiration date or a difference in reactions between Wikipedia and non-Wikipedia firms.

Panel A: Market -adjusted returns (MAR)

Day	Wikipedia (N=330)		non-Wikipedia (N=644)	
	Average MAR (%)	t-stat	Average MAR (%)	t-stat
-10	-0.18	-0.83	0.47	2.67***
-9	0.04	0.18	0.09	0.52
-8	0.17	0.98	-0.18	-1.20
-7	0.16	0.94	0.09	0.57
-6	0.25	-1.47	0.07	0.45
-5	-0.04	-0.21	0.03	0.20
-4	-0.08	-0.45	-0.31	-2.25**
-3	0.13	0.67	0.00	0.02
-2	0.67	3.38***	-0.07	-0.46
-1	0.07	0.40	0.36	2.54**
0	0.14	0.71	0.24	1.37
1	0.08	0.34	0.12	0.67
2	-0.30	-1.70*	-0.41	-3.01***
3	-0.22	-1.27	0.08	0.55
4	-0.00	-0.02	0.13	0.60
5	0.28	1.47	0.20	1.19
6	0.10	0.51	-0.16	-0.96
7	-0.17	-0.86	-0.17	-1.19
8	0.04	0.20	-0.28	-1.75*
9	-0.07	-0.40	0.12	0.84
10	0.44	2.32**	0.11	0.78

Panel B: Cumulative market-adjusted returns (CMAR)

Window	Wikipedia (N=330)		non-Wikipedia (N=644)	
	Average CMAR (%)	t-stat	Average CMAR (%)	t-stat
[-2,+2]	0.65	1.43	0.25	0.69
[-2,-1]	0.74	2.91***	0.29	1.42
[-2,+1]	0.96	2.36**	0.65	1.95*
[0,+2]	-0.09	-0.23	-0.04	-0.15
[-1,+1]	0.29	0.79	0.72	2.32*
[-10,+10]	0.99	1.11	0.55	0.74

TABLE A5 – SENTIMENT ANALYSIS

Table A5 shows sentiment analysis of IPO firm Wikipedia articles and S-1 filings. Panel A displays the average percentage of words in each sentiment dictionary for Wikipedia articles and S-1 filings for 330 U.S. IPOs with Wikipedia articles when going public. LM dictionary is from Loughran and McDonald (2011). Harvard GI dictionary is from the Harvard GI website. Panel B includes OLS regression results for 974 U.S. IPOs from 2006 to 2016. For IPO firms without a Wikipedia article when going public, the sentiment variable is set to zero. The dependent variable *underpricing* is the percentage change from offer price to the first closing price. The Appendix provides variable definitions. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: S-1 and Wikipedia sentiment comparison

<u>Comparison with LM dictionary</u>				
	<u>WIKIPEDIA</u>	<u>S-1</u>	<u>Difference</u>	<u>t-stat</u>
positive	0.71	0.80	-0.09	-2.61***
negative	0.84	1.54	-0.70	-12.95***
net (pos-neg)	-0.13	-0.74	0.61	9.49***
uncertainty	0.36	1.42	-1.06	-41.27***
litigious	0.29	0.97	-0.68	-22.43***

<u>Comparison with Harvard GI dictionary</u>				
	<u>WIKIPEDIA</u>	<u>S-1</u>	<u>Difference</u>	<u>t-stat</u>
positive_GI	1.81	1.79	0.02	0.23
negative_GI	0.61	0.88	-0.27	-6.26***
net_GI	1.20	0.92	0.29	3.28***

Panel B: IPO underpricing and Wikipedia sentiment

VARIABLES	(1) Positive	(2) Negative	(3) Net	(4) Uncertainty	(5) Litigious	(6) Positive_GI	(7) Negative_GI	(8) Net_GI
Sentiment variable	1.909 (2.402)	-0.384 (0.646)	1.363 (1.437)	2.298 (3.055)	-1.801 (1.556)	1.518* (0.835)	-1.306 (0.954)	2.015* (0.958)
VC	8.482** (2.747)	8.459** (2.786)	8.429** (2.797)	8.426** (2.765)	8.434** (2.796)	8.431** (2.675)	8.482** (2.798)	8.431*** (2.652)
top_tier	7.050*** (1.633)	7.280*** (1.434)	7.203*** (1.444)	7.193*** (1.288)	7.297*** (1.434)	6.907*** (1.434)	7.273*** (1.484)	6.825*** (1.434)
overhang	1.518*** (0.327)	1.546*** (0.312)	1.554*** (0.301)	1.511*** (0.296)	1.546*** (0.310)	1.470*** (0.309)	1.544*** (0.310)	1.459*** (0.309)
pos_EPS	4.206 (2.583)	4.189 (2.634)	4.150 (2.648)	4.264 (2.558)	4.108 (2.665)	4.155 (2.526)	4.170 (2.651)	4.084 (2.547)
log_sales	-0.093 (0.612)	-0.058 (0.614)	-0.054 (0.621)	-0.105 (0.632)	-0.046 (0.621)	-0.141 (0.601)	-0.042 (0.618)	-0.127 (0.615)
nasdaq15	0.358 (0.298)	0.347 (0.304)	0.346 (0.305)	0.355 (0.307)	0.350 (0.309)	0.379 (0.299)	0.344 (0.301)	0.380 (0.299)
tech	2.844 (3.023)	3.093 (3.022)	3.106 (3.039)	3.021 (3.020)	3.062 (3.077)	2.806 (2.916)	3.129 (3.031)	2.876 (2.853)
log_age	-1.067 (1.415)	-1.004 (1.397)	-0.932 (1.399)	-1.063 (1.387)	-0.960 (1.388)	-1.165 (1.409)	-0.977 (1.378)	-1.111 (1.398)
log_news	2.732*** (0.846)	2.832*** (0.852)	2.865*** (0.828)	2.737*** (0.828)	2.861*** (0.855)	2.677*** (0.837)	2.853*** (0.844)	2.718*** (0.838)
Constant	-2.951 (3.912)	-3.106 (3.916)	-3.324 (3.839)	-2.879 (3.812)	-3.188 (3.889)	-2.607 (3.862)	-3.096 (3.853)	-2.616 (3.842)
Observations	974	974	974	974	974	974	974	974
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.115	0.114	0.115	0.115	0.115	0.118	0.115	0.121